Time Series Analyses of Global Oil Prices: Shocks, Effects and Predictability

Ruths Sion, Sebastian

(2020)

DOI (TUprints): https://doi.org/10.25534/tuprints-00013411

Publikationstyp: Dissertation

Fachbereich: 01 Fachbereich Rechts- und Wirtschaftswissenschaften

Quelle des Originals: https://tuprints.ulb.tu-darmstadt.de/13411

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Time Series Analyses of Global Oil Prices: Shocks, Effects and Predictability

Dissertation

Approved as fulfilling the requirements for the degree of

doctor rerum politicarum (Dr. rer. pol.)

by the department of Law and Economics
Darmstadt University of Technology

by

Sebastian Ruths Sion, M.Sc.
(Place of birth: Quito)

First supervisor:	**Prof. Dr. Jens Krüger**
Second supervisor:	**Prof. Dr. Michael Neugart**
Filing date:	**29.01.2020**
Date of the viva voce:	**09.07.2020**

Place of publication:	**Darmstadt**
Year of publication	**2020**

D 17

Ruths Sion, Sebastian
Time Series Analyses of Global Oil Prices: Shocks, Effects and Predictability
Darmstadt, Technische Universität Darmstadt
Year thesis published in TUprints 2020
Date of the viva voce 09.07.20

Acknowledgements

This dissertation was written while I was research and teaching assistant at the Chair of Empirical Economics at Darmstadt University of Technology, Germany. Without the support of many people, whom I sincerely thank in the following, the completion of this work wouldn't have been possible.

First and foremost, I am deeply grateful to my supervisor and mentor, Prof. Dr. Jens Krüger. Always open to discussing my work, his knowledge, helpful advice and patience have been invaluable. I would also like to thank Prof. Dr. Michael Neugart for his support as my second supervisor. My interest in economic thinking and research was heavily influenced by his lectures. Additionally, I want to thank Prof. Dr. Dirk Schiereck for chairing the examination committee as well as the committee members Prof. Dr. Volker Nitsch and Prof. Dr. Alexander Kock for their questions and insights during the defense.

Moreover, I am grateful to Dr. Julian LeCrone. The joint work and discussions on early drafts of this dissertation have been very helpful. I thank Sarah Mohring for her productive work and the engaging cooperation. I want to thank Devin Dienes and Yannick Schmidt for their friendship and discussions from a non-economic perspective. I also thank my colleagues in the department of Empirical Economics Vanessa Belew, Dr. Benjamin Hampf, Dr. Mathias Rhiel and Jingwei Pan. Furthermore, my colleagues at Darmstadt University of Technology allowed me to engage in helpful scientific debate and provided me with moral support over the years.

Finally, I am thankful for the continuous support of my friends and my family. I am especially grateful to Esther Klinghammer, for the love, motivation and patience during the final year of writing.

Contents

List of Figures

List of Tables

List of Abbreviations

AIC	Akaike information criterion
AIOC	Anglo-Iranian Oil Company
ARAMCO	Arabian-American Oil Company
ARCH	autoregressive conditional heteroscedastic
bpd	barrel per day
BIC	Bayesian information criterion
CPI	consumer price index
CRSP	Center for Research in Security Prices
DSGE	dynamic stochastic general equilibrium
EIA	United States Energy Information Administration
EJ	Exajoule
EKT	Elliot, Kommunjer, Timmermann
ENET	Elastic Net
FED	Federal Reserve System
FRED	Federal Reserve Economic Database
GDP	gross domestic product
GMM	general method of moments
GNP	gross national product
HAC	heteroscedasticity and autocorrelation consistent
HW	Hamilton & Wu
IIS	impulse-indicator saturation
IOC	international oil company
IPCC	Intergovernmental Panel on Climate Change
IMF	International Monetary Fund
IRENA	International Renewable Energy Agency
IRF	impuls response function
JCPOA	Joint Comprehensive Plan of Action
LARS	least-angle regression

LASSO	least absolute shrinkage and selection operator
LS	least squares
MAE	mean absolute error
mbd	million barrels per day
MCP	minimax concave penalty
ML	maximum likelihood
MSE	mean squared error
MSPE	mean squared percentage error
MZ	Mincer-Zarnowitz
NBER	National Bureau of Economic Research
NIOC	National Iranian Oil Company
NYMEX	New York Mercantile Exchange
OEEC	Organisation for European Economic Cooperation
OECD	Organisation for Economic Cooperation and Development
OLS	ordinary least squares
OPEC	Organisation of Petroleum Exporting Countries
RAP	refiners acquisition price
RMSE	relative mean squared error
SCAD	smoothly clipped absolute deviations
SIS	step-indicator saturation
SVAR	structural vector autoregression
VAR	vector autoregression
SVEC	structural vector error correction
WTI	West Texas Intermediate
WWII	World War II

Chapter 1

Introduction

Of the international events that shaped the global economic order as well as energetic and political interdependencies after World War II (WWII), the oil shocks of 1973/74 and 1979/80 with associated developments around the world, come directly to mind.

Modern life in industrialized societies is surrounded, even more, dependent upon products and services that exist because of crude oil. Oil products fuel engines of cars and trucks required for individual and freight transportation on roads. They also fuel the engines of cargo ships of global maritime trade. Ironically, hydrocarbons also power the giant oil tankers that in 2015 were responsible for the transportation of 61% of global oil flows from the main centers of production to the principal oil consuming and importing regions in Western Europe and East Asia (EIA, 2017). Air transportation also relies primarily on crude oil products such as jet fuel. Other oil products include fuels for generating heat and electricity, asphalt to build roads, as well as petrochemical feedstocks, mainly used to manufacture chemicals, rubber, pharmaceuticals and a large variety of plastics and synthetic materials found in nearly everything we use (EIA, 2019b).

Besides its many uses, the world is drowned in a daily flow of information about oil, in addition to stock market data, exchange rates and gold prices. Oil prices are prominently featured on the front pages of analog and digital business and news media, such as the American Wall Street Journal, the British Financial Times or the German Handelsblatt. Similarly, events that might influence conditions on global oil markets, and thus, global oil prices receive wide coverage, which introduces a second aspect to crude oil that goes beyond its many uses as a commodity or energy source: Supply and demand conditions of oil seem not primarily driven by simple marginal cost and benefit considerations, but also by their concentration in politically unstable regions, such as the Middle East. Due to a strong dependence of Organisation for Economic Cooperation and Development (OECD) countries on oil imports, its analysis has to go beyond simple economics (see e.g. Hamilton, 2011).

The research in this dissertation received during its final stages of editing actuality in the wake of an expected military show-down in the Middle-East between the United States, the largest oil consuming country of the globe and Iran, the second largest oil producing country in the Middle East, the most important oil region in the World.

1.1 Research questions

The dissertation concerns the different aspects of crude oil research, primarily based on four independent empirical analyses, interconnected through a common denominator: time-series analysis methods. It starts with a review of the question, originally at the center of economic research on crude oil: How are macroeconomic performance and oil price movements, especially so-called oil price shocks interrelated? This question, first analyzed empirically by Hamilton (1983) is worth reviewing. While the observation of important economic recessions in the Western World after the oil shocks in 1973/74 and 1979/80 might have been a strong indications of a negative causality in early empirical work, new insights based on longer sample series as well as developments in structural vector autoregression (SVAR) models (see e.g. Kilian, 2009; Kilian and Murphy, 2014) show another picture, which requires differentiation. Based on a broad set of monthly macroeconomic variables for the United States and Germany, the following analysis does not intend to confirm previous results, but to show that different industrialized economies react differently to oil price shocks. This is important as it implies that different energy policies with regard to crude oil might relieve the vulnerability to important oil price movements.

The second empirical analysis concerns the "reversed" oil weapon against important oil exporting countries and its impact on global prices and quantities of oil. The same SVAR models are applied in the framework of the Iran sanctions that were incrementally implemented by the US and the EU from early 2000s onwards. They resulted in a nearly complete boycott of Iranian crude oil exports by 2011, and allow in conjunction with the removal of such sanctions in 2015 and the renewed Iranian crude oil production, to determine their effects on global oil prices. This question is relevant especially in conjunction with the first part on macroeconomic performance and oil price shocks: Is it rational for nations to implement crude oil boycotts if they result in increased global oil prices? Secondly, an analysis of the structural residuals shows that market expectations with regard to future supply disruptions resulting from sanctions, might play a key role in transmitting the effects into crude oil pricing, contrary to price shocks induced by and unexpected supply disruptions.

Departing from the same global model of oil that includes its real price as an endogenous variable, the third analysis is concerned with its oil price forecasting properties. The possibilities to improve forecasting accuracy are explored by applying regularization methods for variable selection. Originating from the "machine learning" literature (see e.g. Murphy, 2012), these methods are now widely used in economic research, especially in cases, where a large number of variables are included in the model. Furthermore, typical lag selection methods, used in the estimation of global models of oil are compared. Finally, the core variable set is augmented by a wide range of possibly relevant regressors as suggested by the literature.

The fourth and final analysis concerns another aspect of oil price forecasting. Based upon the forecasting evaluation framework proposed by Elliott et al. (2005, 2008), that aims at quantifying whether forecasting preferences are asymmetric in a sense that a positive forecast error has a different loss or cost than a negative forecast error of the same magnitude. Originally applied to the forecasts made by individual forecasters or institutions it is proposed as a method to analyze crude oil futures markets. In view of the fact that a futures contract fixes today the quantity and price of crude oil to be delivered in h months, it might be meaningful to view it as the h-step forecast of the spot price of oil. However, the literature suggests that futures remain inaccurate in comparison with forecasts based on economic models or even no-change forecasts (see e.g. Alquist et al., 2013). The common explanation is the risk premium or the expected return that is paid to the long or short side of the contract for taking over the price risk (see e.g. Fama and French, 1987, 1988). Specifically, in the case of oil futures, there is no consensus on how to measure the risk premium, who pays for it and whether it is constant over time (see e.g. Baumeister and Kilian, 2016a). Identifying an asymmetric loss function on market preferences may provide further evidence for the existence of the risk premium and enrich the related literature with further details.

In the following we will present the organization of this dissertation that we follow to tackle the above mentioned research questions.

1.2 Outline of the dissertation

Chapter 2 provides a historical review of the international developments on global oil markets after WWII. The spectacular advancement of crude oil as dominant energy source in industrialized countries in the second half of the 20th century is highlighted. It is important to recognize the strategic role of crude oil during this phase, which can not only be seen as a simple industrial commodity or fossil fuel. Because of the substantial geographical divergence between the main crude oil consuming and crude oil producing regions worldwide, it has to be seen in a broader geopolitical context.

In a similar way, chapter 3 provides an overview on research on oil that was particularly influenced by the oil price shocks in the 70s and 80s. The focus of the first section lies hereby on empirical work regarding the so-called oil price-macroeconomy relationship, first formally observed by Hamilton (1983). The second section is concerned with the theoretical literature which attempts to explain the transmission channels of such an empirically observed relationship. Although both chapters 2 and 3 are not indispensable for understanding the main contents of the study, they provide the context and introduce relevant references.

Chapters 4, 5 and 6 have in common that they rely on global models of oil, first developed by Kilian (2009) and Kilian and Murphy (2014). Chapter 4 formally introduces the models. Departing from a reduced form vector autoregression (VAR) we show why structural assumptions with respect to the reduced form residuals might be required for correct impulse-response analysis, and present some of the identifications methods that exist in this regard. Both global models of crude oil that are used in this dissertation, the recursively identified three-variable SVAR model followed by the sign-restricted four-variable SVAR model are described. After a brief introduction of the core variables, their sources and a discussion of stationarity, the estimation results of both models are presented. The focus lies hereby on impulse-response discussion as well as validation and comparison with the literature.

In Chapter 5 the oil price-macroeconomy relationship for Germany and the United States are reviewed. First, we present the empirical estimation framework, that relies on the estimated structural shocks resulting from the previously described, recursively identified three-variable SVAR model. Furthermore we specify the monthly time series used for Germany and the United States as well as their sources. The empirical results for both countries are then presented and the chapter concludes with a summary and a brief discussion of their implications.

Chapter 6 uses the two SVAR models of chapter 4 to evaluate the impact of the Iran Sanctions in 2011/12 on global oil prices. First, a brief introduction is given to the Iranian oil sector as well as the sanctions imposed by the US and the EU. What follows is an evaluation and comparison of the forecasts of the global oil price implied by both SVAR models. It is based upon the assumption that a forecaster knows the ex-post observed reduction and increase in Iranian crude oil production, induced by the imposition and removal of sanctions in 2011/12 and 2015, respectively. The effect of the Iranian supply disruptions on global oil prices is then estimated using both structural models. A discussion of the observed structural shocks follows, highlighting the possible role of market anticipations in reducing the direct price effects of supply disruptions that can be predicted.

Chapter 7 departs from the reduced form VAR that includes the three core variables previously used in the recursively identified SVAR model and evaluates the forecasts using sparse variable selection methods. First, we introduce the three employed sparse VAR methods, whereas the we discuss the data as well as the transformations that differ from the original work by Kilian (2009). The selection of a benchmark reduced form VAR model follows, before comparing it to the forecasts resulting from regularization methods. The variable set is then enlarged by production indices, exchange rates, investment opportunities and impulse-indicator saturation dummies. We concludes the chapter with a summary and a discussion of the main results.

Chapter 8 applies the forecast evaluation framework, proposed by Elliott et al. (2005, 2008) to crude oil futures in order to offer a different view on their forecasting properties as well as on the risk premium. We first introduce the concept of risk premium on futures contracts and the related literature, before discussing the notion of symmetric and asymmetric loss applied to forecasting. The description of the Mincer-Zarnowitz framework of testing for unbiasedness and efficiency in the case of symmetric loss and serves a discussion of the importance in controlling for the presence of structural breaks within this same framework in the third section. A description of the general method of moments (GMM) estimation of the Elliot, Kommunjer, Timmermann (EKT) loss function that allows for asymmetric loss and how to test for forecast optimality or model validity closes this section. The data, their sources as well as the estimation and test results in case of symmetric and asymmetric loss are then presented. We conclude the chapter with a summary of the results and their implications with regard to futures and their forecasting properties.

Chapter 9 highlights the main results of the empirical analyses and discusses how they can be associated in a common context. The main results are accompanied by outlining possible extensions with regard to crude oil research that is considered as promising. Specific attention is paid to the transition management from oil to renewable forms of energy, required under conditions of climate change.

Chapter 2

Historical developments in oil markets after World War II

> "The numbers - oil production, reserves, consumption - all pointed to one thing: Bigger and bigger scale in every aspect, the oil industry became elephantine." (Yergin, 1991, p. 542)

The following introductory chapter will give a historic overview regarding the evolution of oil markets while concentrating on the post WWII period. Although this work focuses on certain economic aspects of the global oil economy, it is necessary to point towards the complexities around crude oil. In the twentieth century, no other commodity has influenced economics and geopolitics as heavily as crude oil. The overview given will help to get a better understanding of the assumptions that influence this work. Figure 2.1 shows the nominal and real price (in 2010 US$) of crude oil for the North American benchmark West Texas Intermediate (WTI) over the period from January 1950 to December 2017. We include some events that are considered important and typically associated with crude oil. We recommend Yergin (1991) for a detailed and well researched historic recount of the modern oil industry, from it's beginnings in the oil fields of Pennsylvania during the mid 1850s to the end of the twentieth century.

At first glance it becomes apparent that two different oil price regimes seem to have been active between 1950-1974 and from 1974-2017. As Hamilton (2011, p. 8-9) suggests, the importance of the United States as an oil producer and consumer heavily influenced global oil markets in the early post war period. Indeed, until the first oil crisis in the 1970s, global crude oil prices were quoted relatively to the benchmark oil price in the Gulf of Mexico. Thus, the Texas Railroad Commission, the national oil market and price regulator, played an important global role by setting the production quotas an the oil price for the United States.

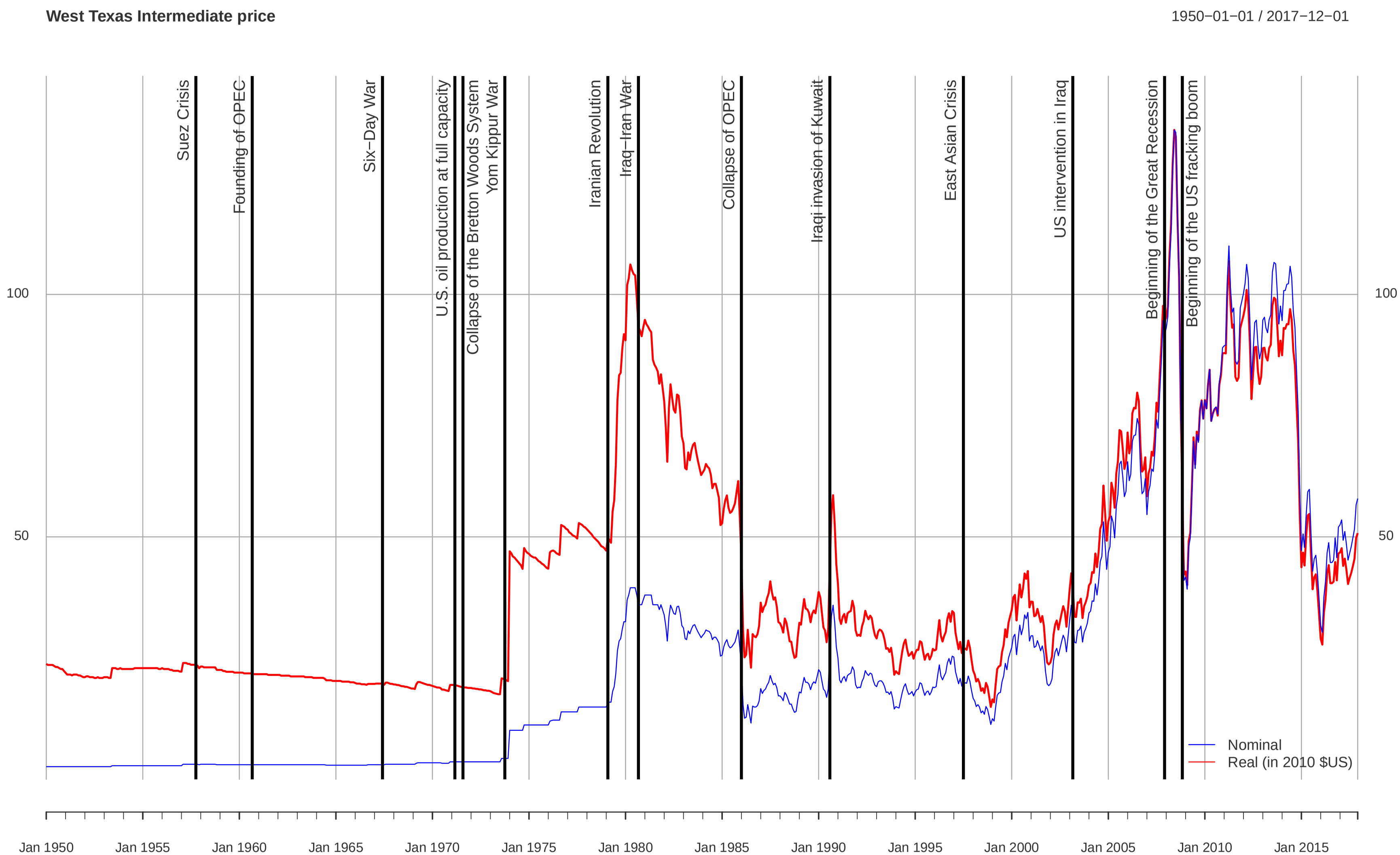

Figure 2.1: Monthly evolution of the nominal and real price of a barrel of WTI crude oil between January 1950 and December 2017 (Source: EIA, FRED).

As we will see in this chapter, the growing import dependence that the United States started to face in the wake of 1973/74 turned the oil markets truly global.

A second aspect to note is the important difference between the price increases and decreases during the 1970s and 1980s and more recent episodes such as the peak in 2008 independent from real or nominal prices. This has to be kept in mind when comparing different episodes in time. We also include some events commonly characterized as oil shocks. Commonly, oil shocks are associated with as an "unexpected oil price increase" (Kilian, 2014, p. 134). Here, oil shocks are viewed from a technical point of view, as the residuals to each variable in a system of equations and may stand for price, demand or supply shocks.

This chapter is organized as follows: Inspired by the evolution of oil prices and associated events, we first describe the organization of oil markets after WWII and introduce important actors such the oil producing companies as well as the main oil importing and oil exporting countries. We then focus on the post 1973 period that is characterized by the oil shocks in 1974/79 and higher market volatility.

2.1 The 50s and 60s: Growing dependence and the rise of Middle Eastern oil

In the same way that coal fueled the Industrial Revolution and thus economic growth in the eighteenth and nineteenth centuries, oil was critical to economic development in the twentieth century. The birth of modern oil industry is attributed to Edwin Drake, who on the 27th of August 1859 successfully drilled for oil near Titusville, Pennsylvania (Yergin, 1991, p.27).[1] It took however over a century for the oil industry to cement it's role as primary energy source in the United States, Western Europe and Japan (Graf, 2014, p. 21).

When looking at numbers the importance of the shift from coal to oil becomes clear. Between 1950 and 1970 global demand for energy grew at a very fast pace. As can be seen in figure 2.2 the primary energy consumption almost tripled from 70 in 1950 to 189 Exajoule (EJ) in 1970.[2] Most noticeable however is that the share of crude oil and related products in total

[1] It should be noted that the first successful drilling for oil in Europe took place in the same year. On the 1st of July 1859, Georg Christian Konrad Hunäus, successfully drilled for crude oil in Wietze, Lower Saxony (Ganser, 2015, p. 42).

[2] 1 Exajoule = 10^{18} Joule

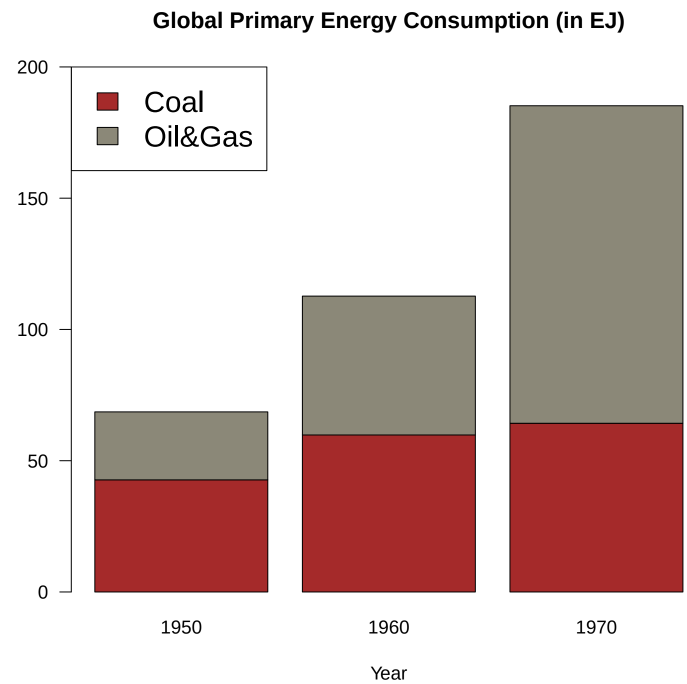

Figure 2.2: World primary energy consumption between 1950 and 1970 (Source: Smil (2000)).

energy consumption increased from 37% in 1950 to 64% in 1970, thus ending more than two centuries of primary reliance on coal. In absolute numbers oil consumption almost increased fivefold from 25.9 EJ to 120.6 EJ over the same period and therefore accounts for most of the additional energy demand.

Main drivers of this spectacular postwar growth were the United States, Western Europe and Japan. As Yergin (1991, p. 541) notes, in the United States daily oil consumption increased from 5.8 million barrel per day (bpd) in 1948 to 16.4 million bpd in 1972. The oil consumption in Western Europe increased from 970,000 to 14.1 million bpd in the same period. Finally, daily oil consumption in Japan increased from 32,000 to impressive 4.4 million bpd.

The special role of the United States and its oil industry in this development is to be briefly highlighted: Until the early 1950s, the United States were responsible for over half of global oil production and consumption (Mayer, 1966, p. 75). Moreover, the US government actively promoted the expansion of an oil-based economy and way of life in Western Europe during the postwar reconstruction. Around 10% of the aid funds granted through the European Recovery Programs, better known as Marshall Plan, were specifically allocated to oil products and financed more than 50% of the oil delivered by American companies to Europe (Graf, 2014, p. 22). During the four years following the hard winter of 1946 and the resulting energy crisis in Europe, it was estimated that up to 20% of the Marshall Plan aid was used for oil imports paid in $ (Yergin, 1991, p. 424). Additionally, the United States Secretary of State George Marshall pushed for the construction of the Trans-Arabian Pipeline in 1950 to transport oil from the newly developed oil fields in the Middle East to Europe (Ganser, 2015, p. 84).

These active American policies in the Middle East have to be seen in a broader strategic effort. After World War II the influence of the former colonial powers was vanishing and the United States saw an opportunity to expand the market shares for its international oil companies (IOCs) through acquisitions of new oil concessions in the Arabian Peninsula. As an early sign of strategic interest in the region, President Roosevelt met in secret with Saudi Arabian King Ibn Saud in February 1945 and reassured the King of American guarantees with regard to the Kingdom's independence and the special diplomatic relations between both countries (Yergin, 1991, p. 404-405). In return King Ibn Saud reaffirmed the oil concession granted to Arabian-American Oil Company (ARAMCO), a joint venture between Standard Oil of New Jersey (to become Exxon) and Socony-Vacumm (to become Mobil Oil), originally granted in 1933 (Yergin, 1991, p. 291). Fearing the British influence in the region however, the King insisted on ARAMCO remaining 100% American (Yergin, 1991, p. 412). Two other major oil concessions: Anglo-Iranian Oil Company (AIOC) (to become British Petrol) operating since 1933 in Iran and the American-Dutch joint venture Gulf-Shell in Kuwait were responsible for the remaining important oil fields in the Middle East. Together, these three major oil deals, were mainly responsible for the extraction and ultimately the transport of vast quantities of oil from the Middle East to Western Europe (Yergin, 1991, p. 422).

The terms agreed between the IOCs and the states of Saudi Arabia, Kuwait and Iran were quite similar. Having originally an average lifespan of 82 years, the concessions covered around 88% of the area of said countries and gave the companies the exclusive right to explore as well as exploit discovered oil reserves. Last, they included complete managerial freedom allowing the companies to take the decisions regarding exploration, development and production unilaterally without having to consult the hosting governments (Stevens, 2012, p. 176-178). Financially, the deals were also quite favorable to the foreign companies. They mostly foresaw upfront payments in form of loans to the host countries and royalties per unit of oil extracted. At a time when the nominal price of a barrel oil was traded at around 2.5$, the royalties paid to the host governments varied considerably. For example Iran received 16.5 cent per barrel from AIOC, while Saudi Arabia received 33 cent per barrel from ARAMCO (Yergin, 1991, p. 444). Faced with costs of production of around 25 cent and costs of distribution of around 50 cent per barrel the oil business in the middle east was extremely profitable to the IOCs (Yergin, 1991, p. 432). Such attractive economic and institutional arrangements in combination with the gigantic proven reserves in the Middle East - in 1968 58.5% of the proved oil reserves were in the Middle East in comparison to 10.1% in North America or 12.1% in the Soviet sphere (Darmstadter, 1971, p. 48) - moved the global center of oil gravity from the Gulf of Mexico to the Middle East.

The seemingly never ending growth of world oil demand and the resulting extraction activities in the Middle East were however destined to change during the 1950s from an institutional point of view. As already mentioned the first concessions were designed so that most of the oil rents went directly to the international oil companies and to a lesser extend the countries in which they paid their income taxes. For example, in 1949 ARAMCO (an American company) paid 43$ million to the United States treasury in income taxes compared to 43$ million in royalties paid to Saudi Arabia (Yergin, 1991, p. 447).

The case of Venezuela, the Latin American country that became the biggest oil exporter (mainly exporting to the United States) and the second biggest oil producer in the 1920s, would become a model to the oil rich countries in the Middle East. Having concessions that were especially beneficial to the IOCs operating in the country, the government saw itself under strong political pressure to renegotiate the concessions originating from the early years of oil development in the country. These renegotiations in 1943 and 1948 resulted in the so called "fifty-fifty" profit split between IOCs and the hosting government (Hults, 2012, p. 422-423). In December 1950 a fifty-fity deal was signed under a lot of pressure between ARAMCO and Saudi Arabia. Similar deals followed soon in neighboring Iraq and Kuwait as the fifty-fifty split of profits became standard practice between IOCs and the hosting countries (Yergin, 1991, p. 447-448).

When the question of renegotiation was raised in Iran in the late 1940s, AIOC the IOC under 51% ownership from United Kingdom, refused to change the terms concerning the legally valid concession of 1933. When in autumn 1950 information of the impending fifty-fifty deal between Saudi Arabia and ARAMCO came to Tehran, the political sentiment became such that an agreement became infeasible. The prime minister Razmara appointed by the Shah to negotiate with AIOC was assassinated in march 1951 (Yergin, 1991, p. 455; Mahdavi, 2012, p. 241). This incident led to the first nationalization of the national oil industry in a middle eastern country and the creation of the National Iranian Oil Company (NIOC) from expropriated AIOC assets (Mahdavi, 2012, p. 241). What followed was an effective embargo by the UK on Iranian oil and the first political and military intervention of foreign powers, namely the United Kingdom and the United States, aiming at replacing the prime Minister Mohammed Mossadegh (Yergin, 1991, p. 464, p. 469-470). By the end of August 1953 the Shah returned to Tehran and appointed a new prime minister in charge of bringing back Iranian oil on world markets. Given the strong reluctance of the IOCs to reengage and resume operations in Iran, it needed pressure from the US and UK governments to convince American, Dutch and French oil companies to enter into a new international consortium with AIOC signaling at the same time the end of the last Sterling-Oil exclusive concession in the world (Yergin, 1991, p. 471-478).

Another important episode in the 1950s, that showed the risks associated with a greater dependence on foreign crude oil, was the Suez Crisis that was triggered on July 26, 1956. During a public speech, Egyptian president Gamal Abdel Nasser, gave the order to Egyptian military forces to seize and gain control of the Suez Canal that connects the Indian Ocean through the Red Sea with the Mediterranean Sea (Yergin, 1991, p. 483). The goal was primarily to gain control of the royalties paid for the passage through the canal in order to finance the construction of the Aswuan Dam on the river Nile (Ganser, 2015, p. 101). Nasser saw this measure as necessary after the United States in accordance with the United Kingdom, withdrew a promised World Bank loan because of Nasser's stance on the state of Israel and his ultimate recognition of "Red China" (Yergin, 1991, p. 482).

The canal was of crucial importance to Western Europe's oil supply. In 1956 around 70% of Western Europe's oil imports, mainly extracted in countries around the Persian Gulf, were transported through the canal to ports in Italy and France (Graf, 2014, p. 53). The expropriation of the Anglo-French Company in charge of running and maintaining the Suez Canal pushed the governments of the United Kingdom, France and Israel to agree on and to elaborate military plans in order to liberate the Canal from Egyptian forces. Plans were ordered into motion on October 29, 1956. Militarily, the operation failed to achieve it's goals, the two former colonial powers and Israel were forced to abandon their intervention after 11 days of combat and agreed to a cease fire on November 6. Under political pressure from both hegemonic powers, the United States and the Soviet Union, the Anglo-French alliance with Israel had to agree to a humiliating retreat in December 1956 (Yergin, 1991, p. 491-492).

From the perspective of oil markets the impact was significant. As Hamilton (2011, p. 10-11) points out, the sinking of 40 ships, effectively blocking the canal through which passed around 1.5 million bpd of crude oil, was very disruptive to oil supplies destined to Western Europe. Additional sabotage actions against the Iraq Petroleum Company's pipeline in Syria, through which around 500 thousand barrel of crude oil were transported to Mediterranean ports, further impacted supplies negatively. The middle eastern crude oil production fell by 1.7 million bpd in November 1956, around 10% of global production. The crisis lasted until April 1957 when most damage on the pipeline through Syria was repaired, the Suez Canal reopened and the nationalization was complete (Yergin, 1991, p. 495).

The crisis, although short lived, had an important impact on the measures that Western European countries would implement in order to better react to possible supply disruptions. The Oil Committee of the Organisation for European Economic Cooperation (OEEC), the European predecessor organization of the OECD, responsible for managing the Marshall Plan aid between European countries, was particularly important in coordinating those efforts (Graf, 2014, p. 54-55). It concluded that shortages as during the Suez episode could repeat themselves and made following five recommendations:

- Creation of national strategic oil reserves corresponding to 80 days of average consumption
- Diversification of suppliers and an increase in the number of oil tankers
- Creation of national emergency committees consisting of representatives from government agencies as well as from the oil industry
- Creation of national emergency plans that would be activated in time of crisis
- Create mechanism of redistribution of reserves and supplies for most hard hit members

Nevertheless, as already shown, crude oil consumption continued to grow as supplies were cheap and abundant, the transition from a coal based to an oil based energy system was, therefore, advancing. This was also true for the United States and fundamentally changed the face of global oil markets. As during the Suez Crisis, the United States has always acted as producer and supplier of last resort (Graf, 2014, p. 54). As can be seen in figure 2.3, US consumption continued to grow while production peaked in the November 1970. In March 1971 the Texas Railroad Commission allowed US production capacity at 100% for the first time since World War II (Dvir and Rogoff, 2009, p. 16). While US demand for oil continued to grow, the slowing national oil production had to be compensated by increasing imports. Between January 1970 and October 1973 alone, imports more than doubled from 1.4 million bpd to 3.4 million bpd. The corresponding import share in total oil consumption increased

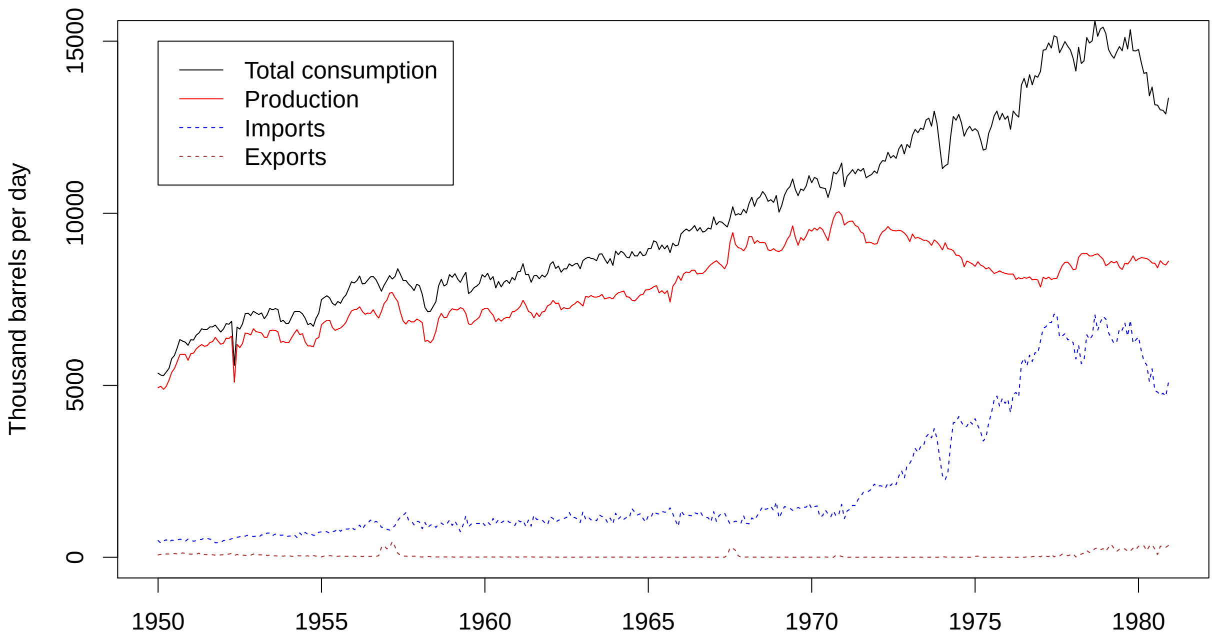

Figure 2.3: Monthly oil consumption, production, imports and exports in the United States between January 1950 and December 1980 (Source: EIA).

from 10% to 28%. The figure also shows that exports have traditionally been very low, with the exception of the episodes during the Suez Crisis in 1957 and the Six-Day War in 1967, when the United States helped containing supply disruptions by increasing their exports.

The receding role of the United States as producer of last resort was followed by demands of renegotiations from Middle Eastern oil producing countries. The formation of the Organisation of Petroleum Exporting Countries (OPEC) in 1960 was the beginning of a shift in power structures. In the early 1970s the Tripoli and Tehran agreements finally abandoned the old system of concessions for a system of participation (Yergin, 1991, p. 582-584). The income share for oil exporting countries was raised from 50% to a minimum of 55%. More importantly, the IOCs no longer owned the oil resources in a given territory, but merely extracted and bought the commodity from sovereign nations. As Dvir and Rogoff (2009, p. 16) remark, the change in ownership rights, gave the oil exporting countries of the Persian Gulf the ability to directly control supply on global markets. Complete nationalization of oil operations and assets would follow in the years to come.

In combination with excess capacity increasingly being built up exlusively in the Middle East, the American delegation to the Oil Committee of the OECD announced to European countries in September 1970 that the spare capacity of the United States was no longer available to allies in case of supply disruptions (Graf, 2014, p. 61). International oil relationships were changing and this change would prove to have important ramifications, when on October 6, 1973 Egyptian and Syrian forces launched a coordinated attack on the state of Israel (Yergin, 1991, p. 588).

2.2 1974 until present: Oil shocks and volatile prices

In retrospect, besides the political goal of Arab oil states to put pressure on western support to Israel, the employment of the "oil weapon" also followed fundamental economic interests. The unilateral termination of the Bretton Woods system by US President Nixon in 1971 cut the revenues of oil exporting countries through devaluation of the US Dollar without a corresponding compensation in the price of crude oil (Hamilton, 2011, p. 13). In early October 1973, members of OPEC met in Vienna with the IOCs in order to renegotiate the posted price of crude oil (Yergin, 1991, p. 599-601). The IOCs offered a 15% increase, around 45 cents more per barrel against the demand from oil exporting countries of three dollars - a 100% increase. After consulting their respective governments, the IOCs announced that such an increase was impossible and declared that negotiations had to be postponed.

This economic aspect that demanded price readjustment was accompanied by difficulties of the oil exporting countries in how to reinvest their oil revenues. In a speech in September 1972, Sheikh Zaki Yamani, the Saudi Arabian oil minister, declared that his country would agree to conclude a strategic partnership with the United States: Saudi Arabia would continue to increase its capacity to extract oil in order to meet the growing demand in the United States. On the other hand, the country expected easier access to investment opportunities in the United States. His strategic offer was ignored however (Graf, 2014, p. 94).

The role of Saudi Arabia was crucial because of its important spare capacity. It could easily overcome shortages induced by other countries by increasing its own production (Graf, 2014, p. 93). Therefore an embargo without Saudi participation was meaningless. . Faced with western support for Israel (particularly from the United States), pressure from other Arab states, the economic consequences of the US Dollar devaluation and the failed attempt to renegotiate with the IOCs through OPEC, the oil producing Gulf states - Saudi Arabia, Kuwait, Abu Dhabi, Iraq, Qatar and Iran - met in Kuwait City on October 16. The countries decided to unilaterally increase their posted price of crude oil from \$2.03 to \$5.11 per barrel, bringing it in line with prevailing prices on the spot market (Yergin, 1991, p. 606). On October 18, the Arab states within OPEC took the decision to cut their monthly oil production by 5% until Israel would withdraw its military forces from occupied Arab territory (Ganser, 2015, p. 183). In November 1973, production from Arab OPEC countries was down 4.4 million barrels per day compared to the level in September, a decrease in 7.5% of global production (Hamilton, 2011, p. 14). As spot market prices continued to increase in December 1973, all OPEC oil ministers met in Tehran to discuss the official OPEC price. Proposals ranged from $8 per barrel, preferred by Saudi Arabia to $23 per barrel, recommended by the Economic Commission of OPEC. The new price set for OPEC oil was agreed to be $11.65. Thus, between October and December 1973, the oil price was submitted to a fourfold increase (Yergin, 1991, p. 625).

A similar episode was to repeat itself a few years later, when protests against the Shah in Iran proved to be persistent and successful in 1978. In the late 1970s Iran was the second largest exporter of crude oil in the world. Daily production averaged around 5.5 million barrel, of which 4.5 million were exported. By November 1978 Iranian oil exports fell to less than 1 million bpd, further destabilizing the country by reducing revenues (Yergin, 1991, p. 678-679). By end of December 1978 oil exports from Iran had virtually ceased altogether as the state descended into chaos. To avoid critical disruptions, other OPEC countries increased their production so that during the first quarter of 1979 global oil supply was only 2 million bpd below the last quarter of 1978. The actual shortage, while small in relation to global consumption of around 50 million bpd, was mainly disruptive as it forced players along the supply chain of oil to make new arrangements in order to replace the lacking deliveries from Iran. In that context, the increasing weight of spot markets was of crucial importance: Buyers payed a premium on every barrel of oil and oil producers saw an opportunity to sell their oil over the official posted OPEC price (Yergin, 1991, p. 685). The spot price of oil increased from around $13 to $34 a barrel. Lastly, fearing a repetition of 1973, the oil companies started to build up oil inventories beyond the real requirements of consumption. This was reinforced by final consumers buying stocks of gasoline, resulting in a self-fulfilling prophecy of additional three million bpd beyond actual global consumption (Yergin, 1991, p. 687).

When Iraqi forces attacked Iran in September 1980 world oil demand was already slowing as the long term response of oil importing countries proved to be persistent (Hamilton, 2011, p. 18). Economic slowdown, energy conservation and significant efficiency gains in western countries were crucial to decrease demand (Yergin, 1991, p. 718). As can be seen in figure 2.4, in 1983 oil consumption in the non-communist world was 45.7 million bpd, around 6 million less than in 1979. On the supply side, the higher oil prices that were set in 1974 in

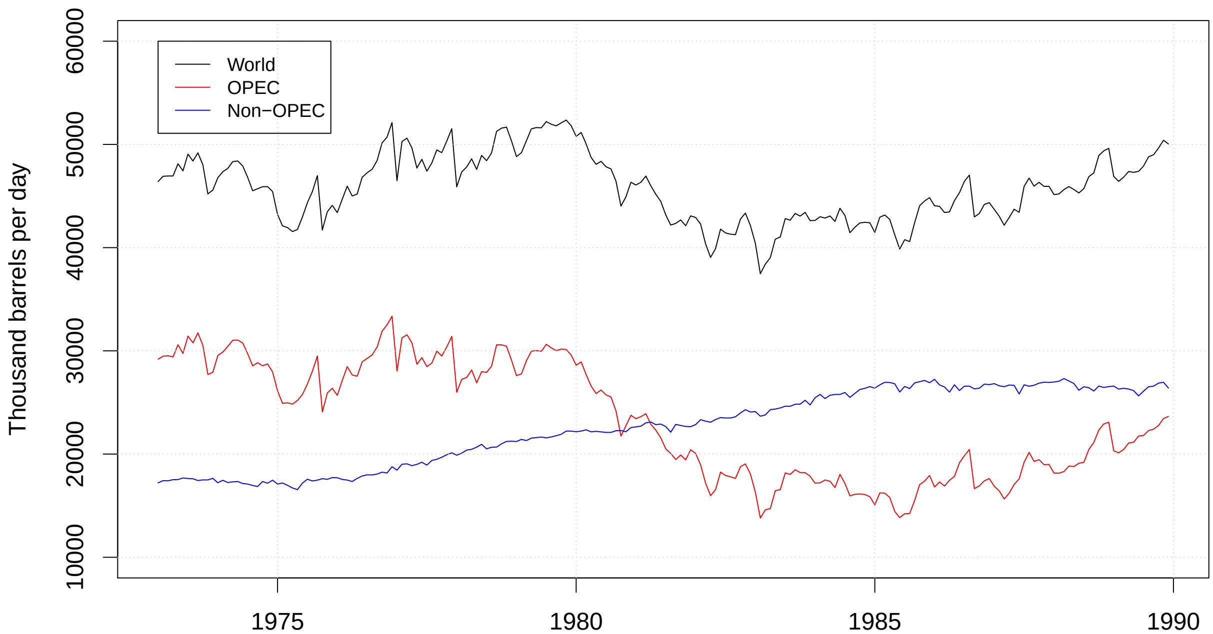

Figure 2.4: Monthly global oil production (excluding the Soviet bloc) between 1973.1 and 1989.12 (Source: EIA).

combination with advances in exploration and extraction technologies, resulted in growing non-OPEC oil supplies. In October 1980, non-OPEC supplies surpassed OPEC supplies and the oil price started to decrease substantially. Saudi Arabia started to cut its production in order to keep the global oil price from falling. In August 1985, with nominal prices around $27 per barrel, OPEC and especially the Saudi Kingdom capitulated, flooding the market with oil, the price fell to less than $10 per barrel in 1986, and the once mighty cartel collapsed (Barsky and Kilian, 2004, p. 131).

The so called collapse of OPEC in 1986 was followed by more than a decade of stable and low oil prices. Figure 2.5 shows the monthly global oil production between January 1990 and December 2017. Worldwide oil production stabilized around 60 million bpd. It should be noted that in comparison to figure 2.4, the production of the former Eastern Block countries is included as they entered world oil markets after the collapse of the Soviet Union. In 1990, the invasion of Kuwait by Iraq and the following military intervention of the United States was accompanied by temporary disruptions in oil supplies. Together, both countries amounted to around 9% of global production (Hamilton, 2011, p. 18). The price of oil briefly doubled from $16 per barrel to $35 one month after the conflict and began to stabilize around $20 per barrel a few months later.

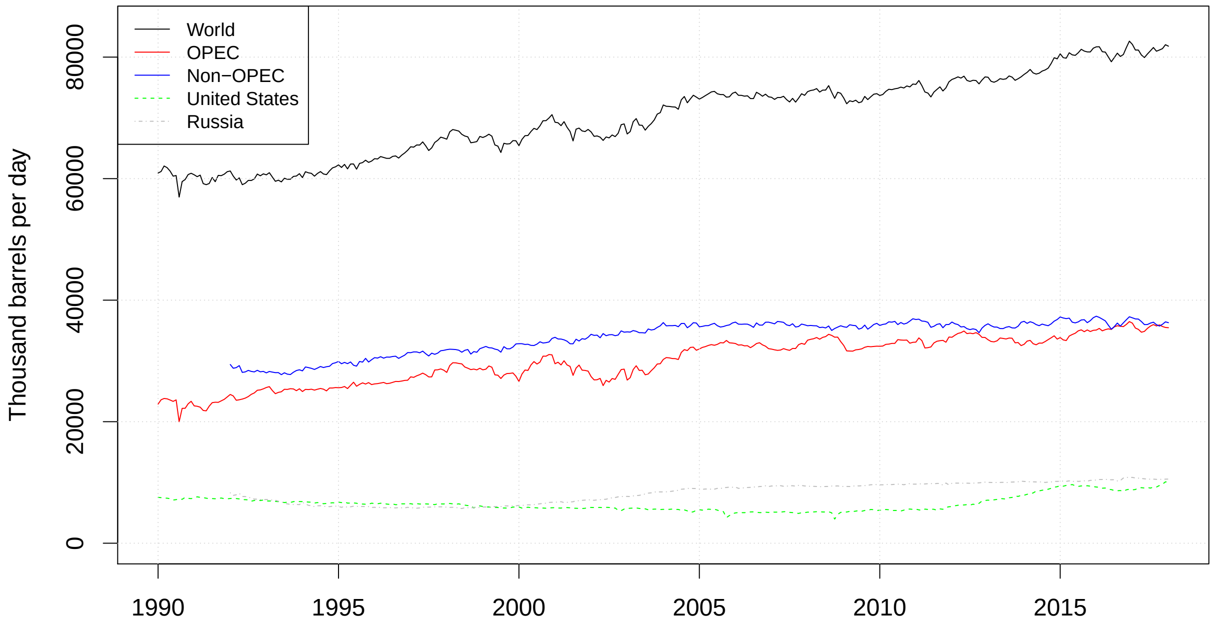

Figure 2.5: Monthly global oil production between January 1990 and December 2017 (Source: EIA).

Starting in 1995, primarily driven by growing the demand in Asia, world oil production increased again. Surplus capacity in combination with falling demand due to the economic crisis engulfing East Asian countries in 1997/98 led to a dramatic fall in prices to as low as \$12 per barrel in December 1998 (Barsky and Kilian, 2004, p. 126). The downturn was however of short duration as the world economy returned to stronger growth rates mainly driven by emerging China. As notes (Hamilton, 2011, p. 21-22) the real world gross domestic product (GDP) grew at an average yearly rate of 4.7% between 2003 and 2005 and at 5% in 2006 and 2007 before the financial crisis of 2008. The resulting increase in oil demand was however not met by significant production increases as figure 2.5 shows. As the short-run price elasticity of oil has been historically low (Hamilton, 2009a, p. 2) a large price increase was necessary to contain demand for oil. The price of a barrel peaked at over \$130 in summer 2008.

Following the global economic slowdown induced by the Great Recession, oil demand fell while overcapacities still persisted. Prices fell to a low of \$32 per barrel in February 2009, but recovered to around \$100 dollars in the first half 2014 as the global economy rebounded and oil demand increased again. This was accompanied by a steady increase of production as can be seen in figure 2.5. (Kilian, 2017, p. 3) notes, that similar to the 1973/74-experience, high oil prices during 2008 created strong financial incentives to explore and expand oil reserves that previously were not economically viable. Accompanied by a fracking boom starting in 2008 in the United States and to a lesser degree in Canada, production also increased in Russia and in some oil producing Gulf States such as Saudi Arabia and Iraq, more than offsetting decreasing output of maturing oil basins, such as the North Sea (Kilian, 2017, p. 3-4). Between mid-2008 and the end of 2017 global oil production increased by around 8 million bpd from 74 to 82 million. The bulk of this increase was driven by the shale oil boom in the United States that almost doubled their production from 5.1 million bpd in summer 2008 to slightly over 10 million in December 2017.

2.3 Summary

This chapter introduced the main characteristics of global oil markets after World War II. The postwar period was defined by a growing demand for oil, mainly driven by the economic recovery of Western Europe and Japan. The United States remained the most important oil producer and consumer giving them a key role in the organization of the postwar oil order. The initial order that relied on a system of concessions and IOCs to extract, transport, distribute and refine oil was being increasingly questioned by the oil exporting countries. More participation and finally nationalization gave members of OPEC more power and influence over production and oil prices. In combination with the growing dependence of the United States on oil imports, the use of the "oil weapon" in 1973/74 was a logical consequence.

As a result oil importing countries reduced their dependence on oil through energy conservation and efficiency measures as well as alternative sources of energy. Economic incentives represented by higher oil prices led to the buildup of further supply capacities on a global scale. In combination, this resulted in more than a decade of very low oil prices that fueled worldwide economic growth in the 1990s and early 2000s. As spare capacity dwindled and supply growth was limited, prices increased to new heights in 2008 reaching over $130 per barrel of oil. The economic downturn induced by the Great Depression also reduced oil demand forcing oil markets down. Recent years have been characterized by worldwide economic recovery and strong growth rates in emerging economies fueling demand for oil. This was accompanied by important supply expansion primarily driven by, but not limited to the United States.

In view of the theoretical and empirical literature review that will follow in the next chapter following historical insights have to be highlighted: Oil price changes were influenced by a number of different factors, individually or in combination. These factors included supply side disruptions such as the closure of the Suez Canal, the 1973/74 oil shock or the recent surge in American shale oil production. On the other hand, there were demand side disruptions such as the decreasing world oil demand in the 1980s or the strong growth of oil demand between 2005 and 2008. However, oil prices were also influenced by factors outside pure supply-demand economics such as colonial and imperial hegemonic interests, inter-cultural and religious conflicts or by the distribution of military and market power between countries.

Chapter 3

Literature review

The important developments during the 1970s and early 1980s, mentioned in chapter 2, attracted the interest of economists, who suspected that the institutional reorganization of global oil markets, growing US-dependence on imported oil, important oil price movements in combination with its poor macroeconomic performance pointed at a natural relationship between oil prices and macroeconomic aggregates (Barsky and Kilian, 2004, p. 115). The resulting literature made the attempt to support this link on theoretical grounds and to search for empirical evidence for such relationship. In this chapter we briefly review main empirical findings in this field of research and provide an overview of the main theoretical explanations.

3.1 Empirical evidence regarding the oil price-macroeconomy relationship

In his seminal work, Hamilton (1983) first remarked, that with one exception, all US-recessions during 1948-1972 were preceded by oil price surges. He made a strong point as to why these oil price shocks were supply driven (typically associated with political events in the Middle East) and thus, regarded oil prices as exogenous to the US economy. In a first step his analysis consists in estimating quarterly bivariate VARs that include the crude oil price and each of the variables proposed by Sims (1980) in his six-variable macroeconomic system: real gross national product (GNP), unemployment, implicit price deflator for non-farm business income, hourly worker wage, import prices and M1 money supply. He then tests for Granger-Causality (Granger, 1969) between oil prices and each of the six variables specifying equation systems, with either four or eight lags.

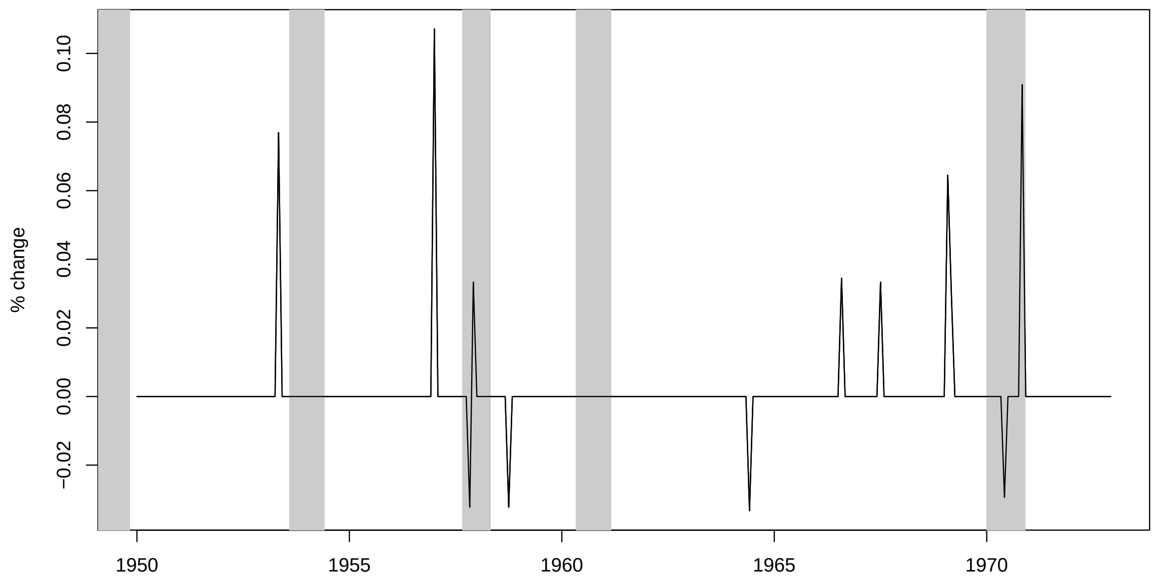

Figure 3.1: Monthly changes in the nominal price of WTI and NBER Business Cycles between January 1950 and December 1972 (Source: EIA, NBER).

His main findings can be summarized as follows: First, oil prices Granger-cause real GNP and unemployment in both specifications. Second, the six-variables do not granger-cause the oil price, indicating that the oil price is exogenous with respect to these variables.[1] These results hold true for different alternative specifications.

Burbidge and Harrison (1984) employ a monthly VAR framework for five different OECD countries and find that oil price shocks had substantial and statistically significant effects on the price level in the US and Canada, while having smaller yet still statistically significant effects on price levels in the UK, Japan and Germany. Looking at industrial production they find significant effects of oil price changes for the US and the UK. These results were complemented by Gisser and Goodwin (1986) who estimated a quarterly St. Louis-type equation expanded by the nominal price of oil over the sample period 1961 to 1982. They find significant effects of crude oil prices on the GNP, the GNP deflator and unemployment in the US. Additional and important contributions that found strong evidence regarding a negative

[1] Only past import prices in the eight lag specification are informative about the price of oil. However, further analysis indicates that it is the component oil import price changes that would not have been predicted given past observations regarding the other variables and that the component of oil price changes that would have been informative about future real GNP and unemployment are not predicted by import price changes (Hamilton, 1983, p. 246).

correlation between macroeconomic aggregates - e.g. output growth or employment - and oil prices include Hamilton (1985) and Santini (1985), who follow similar methodological frameworks.

As Hooker (1996, p. 196) notes, the empirical identification and acceptance of the relationship between oil prices and economic performance lead to widespread inclusion of oil prices as instrumental variables in macroeconomic research. In this regard oil prices have been employed to identify supply and demand in labor markets (Hall, 1991), to diverge from marginal cost pricing (Rotemberg and Woodford, 1996) and to characterize returns to scale (Ramey, 1991). In addition, oil prices have been found to affect the natural rate of unemployment (Carruth et al., 1994), to decrease the impact of technology in business cycle models (Kim and Loungani, 1992), as well as to play a depressing role on investment by affecting uncertainty through price volatility (Ferderer, 1996).

The oil price crash in 1985/86 that followed the period of high oil prices in the 1970s and early 1980s attracted further interest amongst economists. Tatom (1988) was the first to analyze whether the oil price-macroeconomic relationship, identified by the above mentioned studies in the tradition of Hamilton (1983), was also symmetrical. If the effects of oil prices were to be symmetrical, observed decreases should have led to positive effects on macroeconomic aggregates such as GDP growth. His empirical evidence did not reject the so-called null-hypothesis of symmetry, but concluded that effects of oil price increases and decreases are symmetrical on the macro-economic level. Mork (1989, p. 744) on the other hand argued, that the results in Tatom (1988) were influenced by his sample, ending in Q3 1986 and thus not including the important and permanent decline of prices after the collapse of OPEC.

By estimating the same Sims (1980) inspired models as Hamilton (1983) with an expanded sample until 1988:III he found convincing empirical evidence that oil price changes had a weaker influence on macroeconomic aggregates than suggested by the results of Hamilton (1983). By further allowing for oil price increases and oil prices decreases to enter the regression equations individually he quantified both effects. For oil price increases he confirms the results of Hamilton (1983) and for oil price decreases he finds no statistically significant effects concluding that oil prices have non-linear effects on macroeconomic aggregates. Hamilton (2003) further finds strong empirical evidence for linear instability by applying a flexible functional form as developed in Hamilton (2001) to test the null-hypothesis of symmetry against a broad range of consistently estimated alternative non-linear models. He also concludes that oil price increases affect the economy while decreases play no role.

This seemingly time-variant impact of oil prices has also played an important role in research of oil shocks during more recent years. As shown in figure 2.1 on page 8 oil prices started to increase in the late 1990s and even surged after 2004, while US and global economic growth remained strong. Hooker (1996) was the first to question the stability of the relationship between oil prices and the macroeconomy over time and cautioned against using oil prices as instruments in a large body of empirical research for data after 1986. Directly responding to these findings, Hamilton (1996) stressed the fact that one had to take into account that oil price increases simply corrected preceding decreases. He proposed to replace the simple oil price changes series as explanatory variable by the so called "net oil price increase" series. For a given quarter, if the oil price exceeds the previous year's maximum, the percentage change over the previous year's maximum is constructed. If in the current quarter the oil price is lower than it had been at some point in the previous year, the series takes the value of zero. By using this series to reproduce the empirical analysis of Hooker (1996), he confirms the previous negative impact of oil price changes on macroeconomic aggregates.

Within a Philips-curve framework able to accommodate non-linearity as well as structural breaks, Hooker (2002) finds evidence for a structural break on how oil price shocks affect US inflation. While affecting core inflation before 1980, the effects after the break are mostly through the share of oil prices in a price index. Edelstein and Kilian (2007a, 2009) also identify a declining influence of energy price shocks on the US economy. They explicitly attribute this decline to the smaller share of the automobile industry in domestic real GDP. While the effects of increases and decreases on consumption of domestically produced automobiles seem symmetric and stable over time, the overall importance of such changes has declined. Blanchard and Gali (2007) estimate a VAR model that allows for a break in 1980 and further estimate rolling bivariate VAR models that include the price of oil as well as different aggregate variables such as GDP and consumer price index (CPI). They conclude that the US economy had indeed reacted less strongly to oil price changes in the post-1984 sample. They also point at the absence of other adverse shocks, better monetary policy, more flexible labor markets and a smaller share of oil in the economy as factors that recently reduced the vulnerability to oil price shocks. Using the VAR method of Bernanke et al. (1997), Herrera and Pesavento (2009) find that oil price shocks had stronger and longer-lived adverse effects on output growth, the aggregate price level, manufacturing sales growth, and inventory investment during the period preceding Federal Reserve System (FED) chairman Volcker. In addition, their empirical results suggest that monetary policy had a stronger influence on dampening negative effects of oil price changes in the 1970s compared to more recent years.

A common feature of the majority of the studies cited previously is that they take the price of oil or its changes as measure for oil shocks. The common view, heavily influenced by the experiences of 1974 and 1980, was that oil price shocks and associated oil price changes were primarily due to exogenous political events in the Middle East (Hamilton, 1983). However, it is now widely accepted that crude oil prices in particular, are endogenous with respect to US and global macroeconomic developments (Kilian, 2008a, p. 81). In other words, while changes in oil prices might affect macroeconomic aggregates, changes in US and global macroeconomic conditions might in turn also affect oil prices through supply and demand conditions. Not accounting for this endogeneity of the oil price in the estimation of linear regressions, such as those that heavily influenced early research in this field, leads to ordinary least squares (OLS) estimates that are both biased and inconsistent (Davidson and Mackinnon, 2009, p. 311).

One possible solution to the endogeneity problem has been to isolate the exogenous effect of oil shocks, based on the transformation of different variables. A first attempt to formalize this hypothesis was the previously cited "net oil price increase" measure, developed and used by Hamilton (1996, 2003) to explore symmetry and the functional form of the relationship between oil prices and the macroeconomy. Edelstein and Kilian (2007a,b) expand the analysis to series that include net energy price declines, constructed in a comparable way. The use of supply-based variables is a further possibility to account for exogenous oil shocks. Hamilton (2003) for instance created a series that is based on the fall of world oil production attributed to exogenous political events: the Suez Crisis in 1956, the Arab-Israel war in 1973, the Iranian revolution in 1978, the Iran-Iraq war in 1980 and the Persian Gulf war in 1990. The series Q_t takes the value of the largest fall attributed to this event when its start is observed in quarter t and is zero otherwise.

Kilian (2008b) proposes a series for exogenous oil supply disruptions that relies on the construction of a counterfactual world production in the absence exogenous political events affecting oil supply. For a given OPEC country he generates an extrapolated oil production path based on information regarding comparable countries that were not affected by the exogenous event. The difference between actual monthly world production and aggregated monthly OPEC counterfactual oil production is taken as the measure for the exogenous oil supply disruptions. Furthermore Kilian (2008c) assesses the exogeneity of the presented variables using the g_{min} statistic proposed by Cragg and Donald (1993) to identify weak instruments. His test results lead to the conclusion that the presence of weak instruments cannot be ruled out, suggesting that estimates with these variables cause biased coefficient results and hypothesis tests with large errors. He cautions when interpreting results based on those variables.

Different VAR methodologies not only take the endogenous nature of energy prices into account, but also allow to disentangle the underlying causes of energy shocks by imposing structural assumptions. This helps to explain various uncertainties with regard to symmetry and time-variance during the last decades. The observed oil price shocks were combinations of different demand and supply shocks resulting in different outcomes on macroeconomic aggregates (Kilian, 2009; Peersman and van Robays, 2009; Kilian and Murphy, 2014; Kilian and Lee, 2014). As this kind of framework will be used in the next chapters, a more detailed introduction will follow. Before concluding this literature review, we will very briefly go into the theoretical mechanisms that have been proposed to explain the empirical links between oil or energy prices and macroeconomic aggregates.

3.2 Economic theory on oil price shocks

As Kilian (2014, p. 141) points out, economic theory maintained the assumption that oil price shocks were to be considered as exogenous. As previously shown, empirical evidence did not support such view and oil prices are now accepted as endogenous variables. Nevertheless, we will briefly go into the direct and indirect transmission channels of economic theory, under the assumption of true exogeneity. In the standard *supply channel* view, oil is treated as an intermediate input for production and thus oil shocks are viewed as terms of trade shocks as in Kim and Loungani (1992) for example. How crude oil enters the domestic value added production function is "one of the most studied and least resolved issues in empirical macroeconomics" (Backus and Crucini, 2000, p. 196). As imported oil does not enter the production function of domestic value added in a dynamic stochastic general equilibrium (DSGE) framework (Rotemberg and Woodford, 1996, p. 9), per definition oil price shocks can not be seen as productivity shocks with regard to GDP (Barsky and Kilian, 2004, p. 119).

Three proposals have been presented in the literature to deal with this problem by modifying the baseline DSGE of oil importing economies. Rotemberg and Woodford (1996) propose large and time-varying markups to allow for large effects on real GDP by oil price shocks. Atkeson and Kehoe (1999) propose a putty-clay model in which oil price increases are transmitted by capital-energy complementarities to demand for capital services, thus lowering GDP. In the perfect competition model of Finn (2000) energy and oil are essential in order to obtain service flows from capital. The empirical evidence for all three models and by consequence for the domestic supply channel of transmission remains however very weak (Kilian, 2014, p. 142).

The domestic *demand channel* has also been proposed as a direct mechanism of oil price shocks. Edelstein and Kilian (2009) explore how a reduction in discretionary income of households after a rise in energy prices may affect GDP. They conclude, that even with perfectly inelastic energy demand, given the small share of energy expenditure in total consumer spending in the US, the discretionary income effect will be too small to account for large reductions in real GDP. As Hamilton (1988) points out, the discretionary income effect might further be amplified by a rise in operating costs of energy-using durables. Given the possibility of a large dollar value of such goods in relation to the value of energy they use, sectors producing such goods might suffer from the lower demand induced by even small energy price increases. As already mentioned Edelstein and Kilian (2007a) document this effect for the US automobile industry, however, given its small share in aggregate output, its total effect remains small.

Two asymmetric and indirect effects are further mentioned in the literature to explain the impact of changing oil prices on the economy. The *reallocation effect* proposed by Hamilton (1988) arises every time the relative price of oil changes, regardless of the direction of such change. It assumes that there will be inter- or intra-sectoral movements of capital and labor, whenever oil price changes occur. The higher the degree of specialization of an economy, the stronger is the impact on unemployed capital and labor. The second indirect effect refers to the *uncertainty effect* that might arise when oil prices are volatile. Bernanke (1983) argues that firms facing irreversible investment decisions will tend to delay investments when oil prices are volatile. The cash flows from investment projects must depend on the price of oil. Decision-making agents are thus faced with a trade-off between early commitment returns against the benefits of acquiring additional information by waiting. Kilian (2014, p. 144) notes that exactly the same argument holds true for the purchase of energy-intensive consumer durables and that its quantitative importance depends on the relative importance of the price of oil for the investment decision or durable purchase as well as the share of such expenditure in aggregate spending.

Lastly, the relationship between oil price shocks and inflation remains ambiguous: The common view of economists and textbooks postulates that an oil price shock will shift aggregate supply, lowering output and raising inflation (Barsky and Kilian, 2002, p. 137). This perception is at odds, as already mentioned, with very weak empirical evidence for the supply channel of transmission of oil price shocks. As the demand channel seems to dominate in practice, one would expect an oil price shock to shift aggregate demand, lowering both output and prices (Kilian, 2014, p. 141). In this regard, Barsky and Kilian (2002, p. 163) show that oil price shocks are inflationary only for the price of gross output as measured by CPI, but not necessarily for value added as measured by the GDP deflator. An alternative explanation for the apparent interdependencies between oil prices, output and inflation is presented by Bernanke et al. (1997). In their view, it is the reaction of the FED, expecting inflationary pressures from oil price shocks, that amplifies economic contraction by increasing interest rates. If and how central banks should react to oil shocks by adapting mechanically monetary policy remains a further controversial topic in research (Kilian, 2009; Bodenstein and Guerrieri, 2011; Nakov and Pescatori, 2010).

3.3 Summary

This chapter gave a brief overview of the empirical studies and their theoretical explanations regarding the so-called "oil price-macroeconomy" syndrome. Starting with Hamilton (1983) there has been a long tradition of empirical research exploring statistical records for this relationship. As oil markets evolved and data sample sizes increased, new models and specifications were designed to incorporate developments in information previously not accessible. As econometric models and methods evolved the understanding of oil price shocks also increased. Today it is common understanding that oil prices are endogenous to US and global economic developments and are driven by a combination of various supply and demand shocks.

The connection between the channels of transmission as proposed by economic theory and the empirical evidence for the "oil price-macro-economy" relationship found in the data remains controversial. Classical and modified DSGE models view oil price shocks as shifts in aggregate supply reducing output and increasing prices. Empirically, there seems to be more evidence for the demand-side transmission channel that reduces discretionary income or postpones investment and purchase decisions. Again, a key assumption in all theoretical frameworks is that oil prices are strictly exogenous, thus not allowing for feedbacks from aggregate variables to prices.

In the next chapter we will present the advances in VAR and SVAR methodologies before presenting the framework and data that will be used to estimate a global model for oil. The results will subsequently be employed in two studies in chapters 5 and 6.

Chapter 4

Global models of oil: vector autoregression and structural vector autoregression models

Since they were formally proposed by Sims (1980), VAR models and subsequently developed SVAR models remain workhorses in empirical macroeconomic research (Kilian, 2013, p. 515). In this chapter we show how reduced-form VARs have first been used in oil market research and what particularities may arise. We then introduce SVAR models and different estimation methods, followed by the introduction of two SVAR models for the global market of crude oil. After a brief discussion of the data, the chapter closes by showing the empirical results, with a focus on historical structural shocks that have driven oil prices. For an up-to-date review on SVAR models we refer to Kilian and Lütkepohl (2017). The goal of this chapter is to provide an overview of the methodological framework on which the empirical analyses presented in the three chapters to follow. Furthermore it is useful to see what different determinants and developments were historically important to explain oil price movements, thus complementing the context presented in chapter 2.

4.1 Identification methods of structural shocks

Consider the following reduced-form VAR model of order p:

$$\boldsymbol{y}_t = \boldsymbol{\nu} + \boldsymbol{A}_1\boldsymbol{y}_{t-1} + \cdots + \boldsymbol{A}_p\boldsymbol{y}_{t-p} + \boldsymbol{u}_t. \tag{4.1}$$

Standard assumptions apply (Lütkepohl, 2005, p. 69): $\boldsymbol{y}_t$ is a K-dimensional multiple time series with $\boldsymbol{y}_t = (y_{1t}, \ldots, y_{Kt})'$ for a sample period $t = 1, \ldots, T$. Standard estimation methods such as multiple least squares (LS) (Lütkepohl, 2005, pp. 69-82) or maximum likelihood (ML) (Lütkepohl, 2005, p. 87-93) allow consistent estimates of the parameters in $\boldsymbol{\nu}$ and $\boldsymbol{A_i}$, the reduced-form white noise errors $\boldsymbol{u}_t$, and their variance-covariance matrix $\boldsymbol{\Sigma}_u$.

These kind of models were frequently used in the tradition of Hamilton (1983) or Burbidge and Harrison (1984) and included oil prices as well as macroeconomic variables of interest such as the US GDP. Their goal was to estimate impulse response functions (IRFs) to an oil price shock. As already mentioned, because oil prices are endogeneous to global and US macroeconomic aggregates, the components of $\boldsymbol{u}_t$ are likely to be simultaneously correlated resulting in an estimated variance-covariance matrix of $\boldsymbol{\Sigma}_u$ that is not diagonal. The consequences are IRFs that may not reflect the relations between the variables properly (Lütkepohl, 2005, p. 358).

SVAR models have been introduced to deal with this problem allowing to model the instantaneous feedback of the endogenous variables to a so called structural shock. Consider the following general form SVAR model of order p:

$$\boldsymbol{A}\boldsymbol{y}_t = \boldsymbol{\nu}^* + \boldsymbol{A}_1^*\boldsymbol{y}_{t-1} + \cdots + \boldsymbol{A}_p^*\boldsymbol{y}_{t-p} + \boldsymbol{B}\boldsymbol{\varepsilon}_t,$$

where $\boldsymbol{\varepsilon}_t$ stand for the structural shocks that are assumed white noise. We will subsequently focus on the so called B-SVAR model by setting $\boldsymbol{A} = \boldsymbol{I}$. Then the coefficient matrices will be identical to those in the reduced-form. It is common to identify the resulting orthogonal structural innovations $\boldsymbol{\varepsilon}_t$ directly through the reduced-form errors $\boldsymbol{u}_t$ (Lütkepohl, 2005, p. 362):

$$\boldsymbol{u}_t = \boldsymbol{B}\boldsymbol{\varepsilon}_t. \tag{4.2}$$

By inspecting equation 4.2, it becomes apparent, that the reduced-form errors are linear combinations of the structural errors $\boldsymbol{\varepsilon}_t$. The impact multiplier matrix $\boldsymbol{B}$ thus contains the instantaneous impact of the endogenous model variables in period t to a structural shock in the same period. The variance-covariance matrix of the structural terms is typically normalized such that:

$$\boldsymbol{\Sigma}_\varepsilon = \boldsymbol{I}_K. \tag{4.3}$$

To identify $\boldsymbol{B}$ it is helpful to return to the variance-covariance matrix $\boldsymbol{\Sigma}_u$ of the reduced-form errors:

$$E(\boldsymbol{u}_t\boldsymbol{u}_t') = \boldsymbol{B}E(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t')\boldsymbol{B}' = \boldsymbol{B}\boldsymbol{\Sigma}_\varepsilon\boldsymbol{B}' = \boldsymbol{B}\boldsymbol{B}' \tag{4.4}$$

where use is made of $\boldsymbol{\Sigma}_\varepsilon = I_K$ in the last step. As already indicated, the reduced-form variance-covariance matrix $\boldsymbol{\Sigma}_u$ can be estimated consistently and is thus known. If one considers $\boldsymbol{\Sigma}_u = \boldsymbol{B}\boldsymbol{B}'$ as a system of linear equations for determining the unknown parameters in $\boldsymbol{B}$. Because of the symmetrical nature of variance-covariance matrices, $K(K+1)/2$ different equations with $K(K-1)/2$ restrictions suffice to identify all K^2 elements of $\boldsymbol{B}$.

A common and simple approach has been the identification of $\boldsymbol{B}$ by a Choleski decomposition, meaning that by construction $\boldsymbol{B}$ is lower triangular (Kilian, 2013, p. 518). The mechanical application of such recursive identification strategies was widespread in the 1980s but was soon heavily criticized (see Cooley and LeRoy, 1985, for example). As Kilian (2013, pp. 519-522) notes, such atheoretical models made strong assumptions by imposing a particular causal chain rather than observing the causality in the data. Notably, arranging the K variables differently in the VAR model will result in different solutions for the matrix $\boldsymbol{B}$. Thus, without assumptions based on economic theory about the recursive order imposed, the solution may have no plausible economic interpretation. In section 4.2.1, we will present the first global SVAR model for oil and discuss the underlying economic assumptions that allow us to justify a specific arrangement of the variables in the context of a 3-variable SVAR model.

The critique towards atheoretical recursively identified models led to the development of models which rely on alternative non-recursive short-run restrictions in the matrix $\boldsymbol{B}$ based on economic theory (see Bernanke, 1986; Sims, 1986; Blanchard and Perotti, 2002, for example). Another methodological approach for the identification of $\boldsymbol{B}$ imposes restrictions on the estimated IRF of a variable given a certain structural shock (see, for example, Blanchard and Quah, 1989). Such an approach is theoretically motivated by short-run versus long-run relationships. For example, most economists would agree that demand shocks resulting from monetary policy are neutral in the long run, while shocks in productivity are not (Kilian, 2013, p. 529).

A further identifying approach that was proposed by Faust (1998), Canova and Nicolo (2002) and Uhlig (2005) relies on sign restrictions in the impact multiplier matrix $\boldsymbol{B}$. These are increasingly popular amongst researchers, as sign restrictions can usually be directly supported by economic theory. As Kilian (2013, pp. 534-538) remarks that sign restrictions are not without drawbacks in comparison to other approaches. First, there is a misconception amongst users that sign restricted SVAR models are more general and thus more credible than alternative SVAR models, which is not true. Second and more importantly, there is no unique solution to sign restricted SVAR models. We rather identify a set $\boldsymbol{B}^*$ of admissible solutions $\boldsymbol{B}$ to equation 4.4. Given that each admissible solution in $\boldsymbol{B}^*$ is likely to result in different IRFs, the choice of the most likely solution remains, therefore, an ongoing research question.

As we will see for the case of a 4-variable global SVAR model for oil, imposing further restrictions may be a solution to narrow down the number of admissible models considerably (Kilian and Murphy, 2012, 2014). Indeed, by considering additional restrictions such as lower and upper bounds for demand or supply elasticities as suggested in other empirical studies, or long-term IRF restriction based on theoretical grounds, it is possible to obtain solutions that are narrower.

4.2 Model specifications for the global crude oil market

In this section we will present, estimate and compare the two SVAR models that we will later apply to two empirical exercises in chapters 5 and 6. First, we present the 3-variable recursively identified model proposed by Kilian (2009). We then introduce the sign-restriction identified 4-variable model as in Kilian and Murphy (2014). In the latter case we will also describe the numerical algorithm to identify the admissible solution set $\boldsymbol{B}^*$.

It is useful to first specify the reduced-form of the monthly VAR(p) for K variables:

$$\boldsymbol{y}_t = \Sigma_{i=1}^{24} \mathbf{A}_i \boldsymbol{y}_{t-i} + \mathbf{u}_t. \tag{4.5}$$

The corresponding impact response relation to structural shocks is again defined as:

$$\mathbf{u}_t = \mathbf{B}\boldsymbol{\varepsilon}_t. \tag{4.6}$$

We adopt the common practice for monthly VARs in setting the order to $p = 24$ in the following. As Kilian and Lütkepohl (2017, p. 54) note, this permits to correctly capture delayed responses to shocks from variables in the system. To control for seasonal monthly effects seasonal dummies are included. However, these are suppressed for notational convenience.

4.2.1 Recursively identified model of the global market for crude oil

To address the issue of reverse causality from macroeconomic aggregates to oil prices, as well as to disentangle what shocks drive the price of oil on a global scale, Kilian (2009) developed an SVAR model with three variables. He proposed $\boldsymbol{y}_t = (\Delta prod_t, rea_t, rpo_t)'$ as endogenous variables, where $\Delta prod_t$ is the percentage change in global crude oil production, rea_t denotes the index of real economic activity and rpo_t refers to the real price of oil. Data sources and specifications are analyzed and discussed in more detail in subsection 4.3.

As for the structural errors in ε_t, he builds on the earlier insights of Barsky and Kilian (2002, 2004) and distinguishes two demand and one supply shock: *flow supply shocks* (shocks to the physical supply of crude oil), *flow demand shocks* (shocks to the demand for crude oil due to variation in the global business cycle) and *oil-specific demand shocks* (shocks due to all other factors such as precautionary demand, weather effects or technology shocks). Again, we keep in mind that all structural shocks are orthogonal so that effects not captured by the first two shocks are captured by the third one. The following recursive structure is postulated for the 3×3 matrix $\boldsymbol{B}$ in equation (4.6):

$$\begin{pmatrix} u_t^{\Delta prod} \\ u_t^{rea} \\ u_t^{rpo} \end{pmatrix} = \begin{bmatrix} b_{11} & 0 & 0 \\ b_{21} & b_{22} & 0 \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{pmatrix} \varepsilon_t^{\text{flow supply shock}} \\ \varepsilon_t^{\text{flow demand shock}} \\ \varepsilon_t^{\text{oil specific demand shock}} \end{pmatrix} \tag{4.7}$$

The restrictions included in the matrix $\boldsymbol{B}$ in equation (4.7) are motivated as follows (Kilian, 2009, p. 1059). The model assumes a vertical oil supply curve as can be seen in figure 4.1 where P is the price of crude oil and Q the quantity of crude oil available on the market. The shifts resulting from the structural flow supply shock and the two structural demand shocks in the figure are based on following theoretical considerations:

- Crude oil supply shocks ($\varepsilon_t^{\text{flow supply shock}}$) are the unpredicted innovations to global crude oil production ($u_t^{\Delta prod}$). Standing for the classical notion of an oil shock, they include supply disruptions due to political events in oil-rich countries or reductions on the basis of decisions taken by OPEC. Global crude oil production is assumed not to respond to either of the two demand shocks. Such assumption seems plausible as oil-producing countries will be sluggish to costly readjustments of their production when faced with uncertainty about the longevity of demand shifts.

- Unexpected shifts in the global real economic activity (u_t^{rea}) that are not explained by flow supply shocks, will be specified as shocks to the international demand for industrial commodities (including crude oil) due to the global business cycle ($\varepsilon_t^{\text{flow demand shock}}$). Oil-specific demand shocks are assumed to affect real economic activity with a delay, excluding an instantaneous feedback. This is consistent with the delayed economic downturns (recessions occurring some quarters after the oil price hikes) as discussed

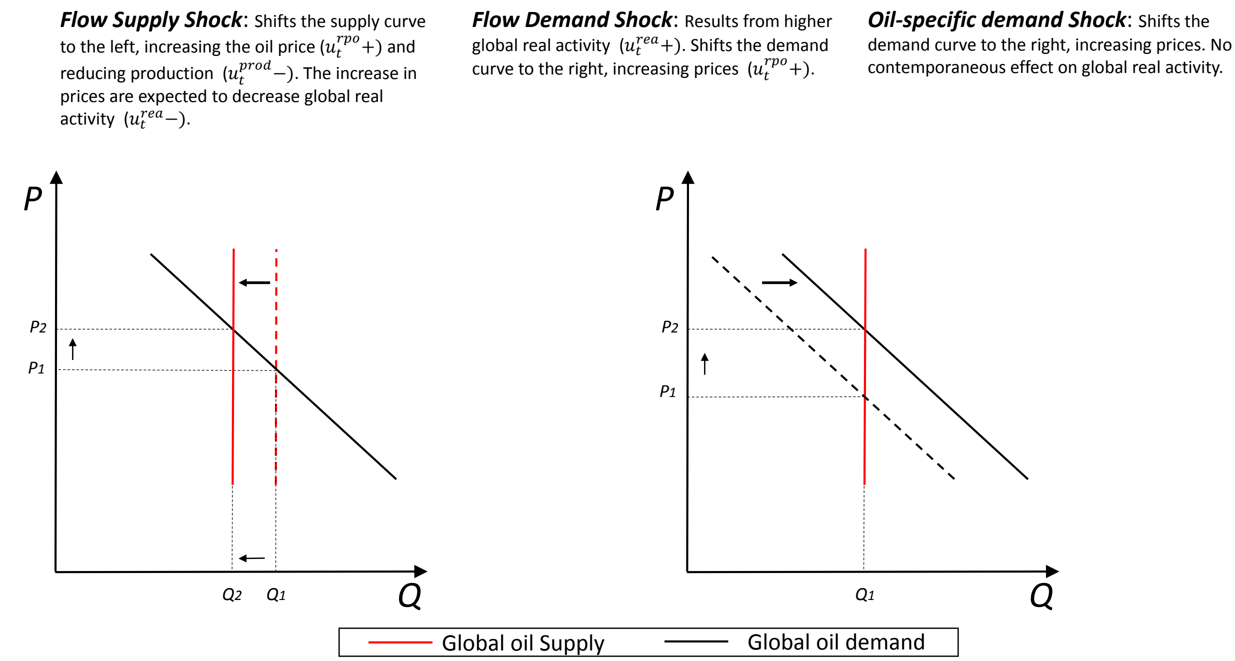

Figure 4.1: Structural shocks in the Kilian (2009) SVAR model.

following major oil episodes in chapter 2.

- Finally, oil-specific demand shocks ($\varepsilon_t^{\text{oil specific demand shock}}$) are defined as innovations to the real price of oil (u_t^{rpo}) that can neither be explained on the basis of oil supply shocks nor by demand shocks for oil resulting from shifts in global demand for industrial commodities. As can be seen, this oil-specific shock will in particular capture shifts in precautionary demand resulting from uncertainty about future oil supply, a behavior consistent with theoretical models about expectations (see, for example, Alquist and Kilian, 2010).

From equation (4.7) and figure 4.1 it becomes clear, that some directional assumptions are based on theoretical grounds. Disruptions in oil supplies (a negative flow supply shock) will result in oil price increases, while oil price increases will disrupt global real activity. However, these sign assumptions are not imposed on the model as all the non-zero coefficients in the matrix $\boldsymbol{B}$ can result in a negative or positive estimate. The next model does not require exclusion restrictions but will impose sign restrictions on the coefficients in $\boldsymbol{B}$.

4.2.2 Sign restricted model of the global market for oil

Building on the insights of the recursively identified model of the global oil market, Kilian and Murphy (2012) first proposed a sign-restricted approach to identify $\boldsymbol{B}$ in the three-variable case. In Kilian and Murphy (2014) they added a fourth endogenous variable in the form of crude oil inventories to further disentangle oil specific demand shocks into precautionary demand and other demand for oil in Kilian and Murphy (2014). Their goal is to explicitly model speculation on the oil market, which is defined from an economic point of view as a purchase with no intention of current consumption. Usually speculation is the result of anticipated increases in future oil prices (Kilian and Murphy, 2014, p. 455). They propose $\boldsymbol{y}_t = (\Delta prod_t, rea_t, rpo_t, inv_t)'$ with the additional variable being month to month absolute changes in above-ground oil inventories, foreseen by OECD, as a proxy for global oil inventory changes. The following sign restrictions are postulated for the 4×4 matrix $\boldsymbol{B}$ in equation 4.6:

$$\begin{pmatrix} u_t^{\Delta prod} \\ u_t^{rea} \\ u_t^{rpo} \\ u_t^{inv} \end{pmatrix} = \begin{bmatrix} - & + & + & . \\ - & + & - & . \\ + & + & + & . \\ . & . & + & . \end{bmatrix} \begin{pmatrix} \varepsilon_t^{\text{flow supply shock}} \\ \varepsilon_t^{\text{flow demand shock}} \\ \varepsilon_t^{\text{speculative demand shock}} \\ \varepsilon_t^{\text{other oil-specific demand shock}} \end{pmatrix} \tag{4.8}$$

Instead of three, four structural shocks are explicitly taken into consideration. As in the first model, we include classical crude oil supply shocks ($\varepsilon_t^{\text{flow supply shock}}$) and demand shocks for oil due to the global business cycle ($\varepsilon_t^{\text{flow demand shock}}$). Shifts in demand for above-ground oil inventories are captured by the third structural shock ($\varepsilon_t^{\text{speculative demand shock}}$). The fourth shock ($\varepsilon_t^{\text{other oil specific demand shock}}$) captures all other idiosyncratic demand effects not accounted for by the first three structural shocks. We will now give a theoretical explanation for the sign restrictions for $\boldsymbol{B}$ in equation 4.8 based on Kilian and Murphy (2014, p. 458). Figure 4.2 presents the mechanism based on the global supply and demand curves for oil. First we should note that the structural shocks have been normalized so that they all induce an increase in the real price of oil, as can be seen in the 2nd and 3rd column of third row of the $\boldsymbol{B}$ matrix. Secondly, Kilian and Murphy (2014) allow for changes in crude oil supply within a month of a given shock, so that the oil supply curve is no longer vertical. The specific reaction of all endogenous variables to each structural shock are now presented.

- Given a negative flow supply shock, defined as shift to the left of the oil supply curve, oil production will decrease, the real price of oil and by consequence real economic activity will increase in the same month. In the case of oil inventories, two effects should

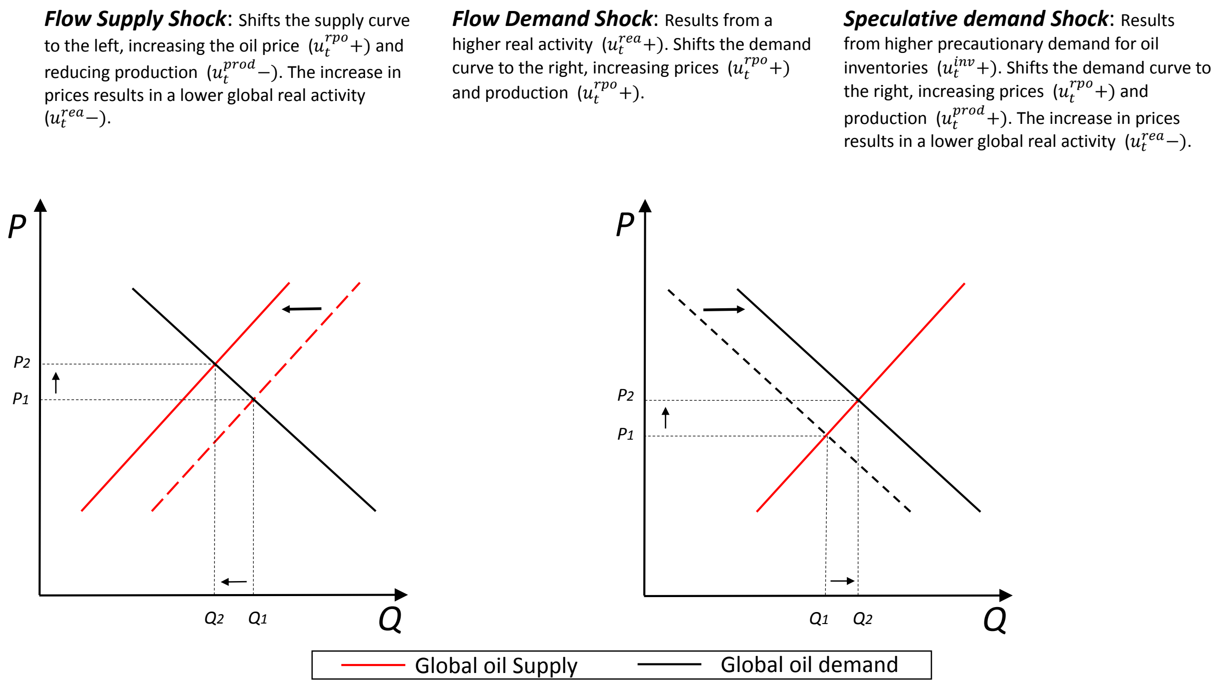

Figure 4.2: Structural shocks in the Kilian and Murphy (2014) SVAR model.

occur simultaneously: On the one hand, inventories should decrease in an attempt to compensate for the loss of supplies. On the other hand, as oil prices increase, the precautionary effect will induce the accumulation of inventories. It remains to be seen which effect dominates in practice, so no sign restriction is imposed.

- A positive flow demand shock resulting from a shock to global demand for industrial commodities, so that we observe increasing global real activity is associated with a shift to the right of the contemporaneous oil demand curve. It will increase global oil production and increase the real price of oil. As in the case of a negative flow supply shock, the effect on inventories remains uncertain, so that again no sign restriction is imposed.

- Speculative demand shocks, again associated with a shift to the right of the contemporaneous oil demand curve, will instantaneously result in a higher oil production and oil prices. In comparison to flow demand shocks, the price hike is not accompanied by strong global demand for industrial commodities, thus we expect the impact on the real activity index to be negative. As the increase in demand is due to precautionary pressures with respect to future oil price increases or supply flow disruptions, an accumulation of crude oil inventories can be observed.

- Other oil-specific demand shocks are solely restricted in the sense that they are also associated with oil price increases. As their effects on the other endogenous variables are difficult to predict based on theoretical grounds, no other sign restrictions are imposed.

From both equations (4.7) and (4.8) it becomes apparent why historically observed oil price hikes were different in their resulting effects on macroeconomic aggregates. Thus, when observing oil price increases without modeling the underlying causes correctly, we might find conflicting results for crude oil price increases of the same magnitude. For example, if an oil price shock is primarily the result of a flow demand shock resulting from strong global activity, the observed effects might even be positive on US economic growth in comparison to oil price increases resulting from a drop in crude oil supply.

Additional impact and dynamic restrictions

Relying uniquely on sign restrictions can lead to IRFs that are quite different. Kilian and Murphy (2012) thus propose to include all available information about the market structure to reduce the admissible set of solutions to a smaller number of qualitatively comparable estimates. Let hereby $\tilde{\boldsymbol{B}}$ be an estimate of $\boldsymbol{B}$ belonging to the solution set. Kilian and Murphy (2014, pp. 461-462) propose the following additional market structure restrictions that must be true in global oil markets:

- Based on the work done by Hamilton (2009b) and Kellog (2011) that postulate that oil supply elasticity is very low if not essentially zero (as implied by the vertical supply curve in the recursively identified SVAR model), they impose an upper bound of 0.025 on the *impact elasticity of oil supply*. A direct estimate of this measure may be constructed from the coefficients $\tilde{b}$ in the matrix $\tilde{\boldsymbol{B}}$, where $\tilde{b}_{ij}$ is the coefficient in the i-th row and j-th column, by evaluating the ratio of the impact responses of oil production and of the real price of oil to an increase in flow demand or speculative demand:

$$0 \leq \frac{\tilde{b}_{12}}{\tilde{b}_{32}} \leq 0.025 \quad \& \quad 0 \leq \frac{\tilde{b}_{13}}{\tilde{b}_{33}} \leq 0.025$$

- A further bound restriction that can be applied to an estimate of the *impact elasticity of oil demand in use*. A lower bound is motivated based on the estimates of Hausman and Newey (1995) that consistently find US household long-run price elasticities of around -0.8. The upper bound is given by zero as the elasticity is expected to be slightly negative. Given oil production Q in period $t-1$, Q_{t-1}, and average changes in oil inventories over the sample period given by $\overline{\Delta S}$, demand elasticity can be estimated as follows:[1]

$$\eta_t^{use} = \frac{\frac{(Q_{t-1} \times \tilde{b}_{11}/100) - \tilde{b}_{41}}{Q_{t-1} - \overline{\Delta S}}}{\tilde{b}_{31}/100}$$

As this impact elasticity in use is time-dependent, the average oil demand elasticity in use, denoted $\bar{\eta}^{use}$, over the sample period can be bounded as follows:

$$-0.8 \leq \bar{\eta}^{use} < 0$$

- Lastly, they impose restrictions on the *cumulative impulse responses to a flow supply shock*. *Oil production* and *real activity* respond negatively in the first 12 months after an oil supply disruption. The *real price of oil* is restricted to react positively for the 12 first months after a negative flow supply shock. These restrictions are necessary to rule out models that result in price reduction when such a negative shock occurs.

Identification strategy

Given the reduced-form VAR estimate of the variance-covariance matrix $\widehat{\boldsymbol{\Sigma}}_u$ we construct the eigendecompostition $\widehat{\boldsymbol{\Sigma}}_u = \boldsymbol{P \Lambda P'}$ and define $\boldsymbol{B} = \boldsymbol{P\Lambda}^{0.5}$ such that $\boldsymbol{B}$ satisfies $\widehat{\boldsymbol{\Sigma}}_u = \boldsymbol{BB'}$. Then for any orthogonal $N \times N$ matrix $\boldsymbol{D}$, $\tilde{\boldsymbol{B}} = \boldsymbol{BD}$ also satisfies condition (4.4), i.e. $\tilde{\boldsymbol{B}}\tilde{\boldsymbol{B}}' = \widehat{\boldsymbol{\Sigma}_u}$.

Thus, following Rubio-Ramirez et al. (2010) we create a set $\tilde{\boldsymbol{B}}^*$ of admissible solutions by applying following loop with a large number of draws (Kilian and Murphy, 2014, p. 463):

1. Draw an 4×4 matrix $\boldsymbol{K}$ of NID(0,1) random variables. Obtain the $\boldsymbol{QR}$ decomposition of $\boldsymbol{K}$ so that $\boldsymbol{K} = \boldsymbol{QR}$ with $\boldsymbol{QQ'} = \boldsymbol{I}_4$ and $\boldsymbol{R}$ being upper diagonal.

2. Define $\boldsymbol{D} = \boldsymbol{Q'}$ and construct the structural impact response matrix $\tilde{\boldsymbol{B}} = \boldsymbol{BD}$ to compute the implied structural impulse responses. If all restrictions are satisfied retain $\tilde{\boldsymbol{B}}$ else discard $\tilde{\boldsymbol{B}}$.

In the next section we will discuss the data that will be used to estimate both SVAR models presented above for the global market of crude oil.

[1] We refer to Kilian and Murphy (2014, pp. 477-478) for a formal discussion of how to construct this elasticity from the structural VAR model.

4.3 Data

As pointed out the above mentioned models rely on three monthly core variables: global crude oil production, the real activity index as a measure of global demand for industrial commodities and a measure for the global price of oil. In the second model we also use data on above-ground crude oil inventories. In the following we will briefly discuss all four time series, their source as well as the potential transformations needed in order to ensure stationarity. Our sample period extends from January 1974 to December 2017, implying a sample size $T = 528$. While all three studies start in the same period, we note that this study extends the sample periods of Kilian (2009) and Kilian and Murphy (2014) significantly. The former ends in December 2007 and the latter in December 2009. The sample selected in this study thus covers the important global developments during the Great Recession as well as the recovery thereafter.

4.3.1 Crude oil production

The first core variable is the global crude oil production (figure 4.3). The raw data is retrieved from the *Monthly Energy Review* table 11.1b as published by United States Energy Information Administration (EIA) and is expressed in thousand barrels per day.[2] As we already discussed the historical developments that drove global oil production in our sample in chapter 2, here the discussion is limited to the characteristics of the time-series. As can be seen in the first row in figure 4.3 the series shows a clear trend pointing visually to unit-root nonstationarity. Kilian (2009) and Kilian and Murphy (2014) propose to include the series as percent changes into the model (expressed as differences in logarithms). The variable will be hence called $\Delta prod_t$. Visually, the time series shows a decrease in volatility while appearing stationary in the mean.

We test the unit-root null hypothesis using the ERS test of Elliott et al. (1996). Additionally, we test the null hypothesis of stationarity using the KPSS test of Kwiatkowski et al. (1992) to further validate the ERS test results. The outcome of both tests for the $\Delta prod_t$ series can be seen in table 4.1. The first and second column contain the test statistics for the ERS test given a model with constant and trend respectively. We further include the significance levels at which the null hypothesis of stationarity can be rejected in both columns. The third and fourth column contain the test statistics for the KPSS test given a model with constant and trend respectively. Again, we include the significance levels at which the null hypothesis of

[2] The *Monthly Energy Review* is a monthly updated publication of historical and recent energy statistics that can be accessed through https://www.eia.gov/totalenergy/data/monthly.

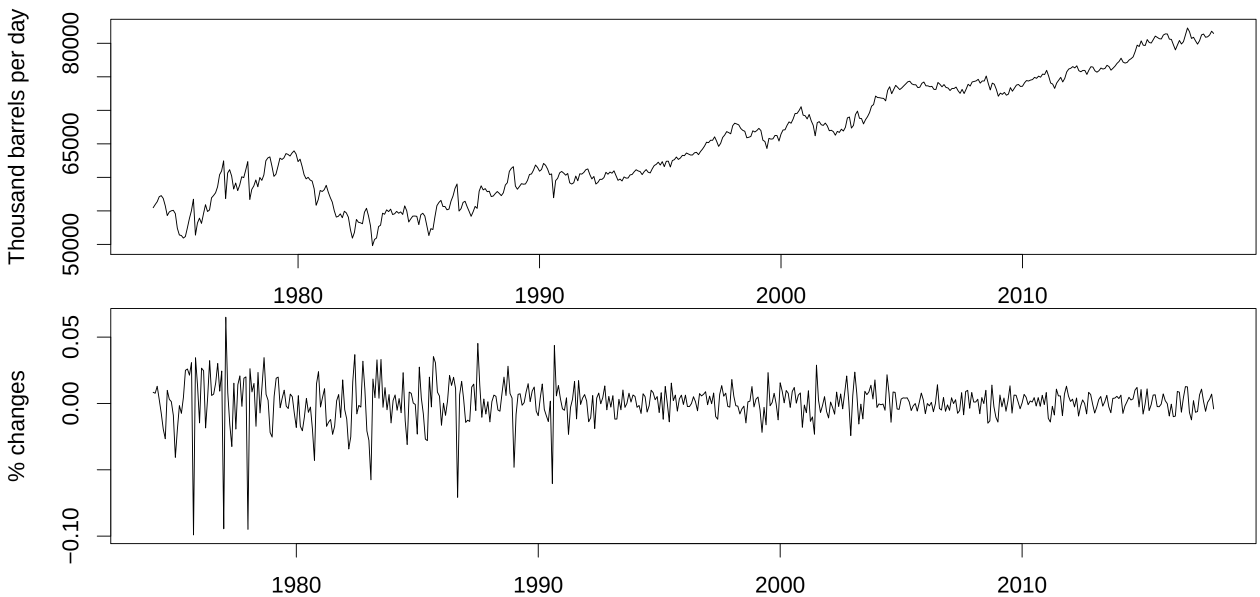

Figure 4.3: Global oil production in thousand barrels per day and monthly % changes (Source: EIA).

Lag order	ERS test (H_0: unit-root)				KPSS test (H_0: stationarity)	
	Model with constant		Model with trend		Model with constant	Model with trend
0	-20.747	***	-23.462	***	0.031	0.016
1	-13.861	***	-16.547	***	0.034	0.017
2	-10.829	***	-13.596	***	0.037	0.019
3	-8.685	***	-11.394	***	0.041	0.021
4	-7.778	***	-10.670	***	0.044	0.022
5	-6.117	***	-8.687	***	0.050	0.025
6	-5.644	***	-8.288	***	0.053	0.027
7	-4.789	***	-7.226	***	0.058	0.029
8	-3.784	***	-5.829	***	0.061	0.031
9	-3.437	***	-5.392	***	0.062	0.031
10	-3.368	***	-5.378	***	0.063	0.032
11	-2.770	***	-4.537	***	0.065	0.033
12	-2.789	***	-4.636	***	0.064	0.032
Significance:	***=1%,	**=5%,	*=10%			

Table 4.1: Stationarity evaluation of the $\Delta prod_t$ series.

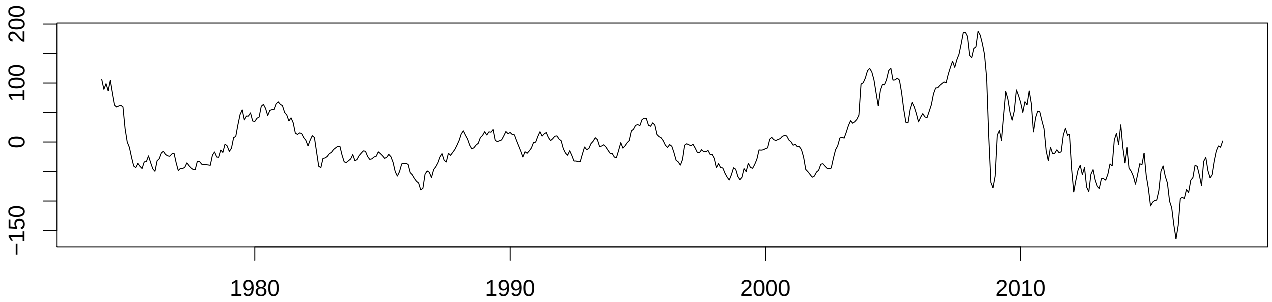

Figure 4.4: Real activity index (Source: https://sites.google.com/site/lkilian2019/research/data-sets).

stationarity can be rejected. In the case of $\Delta prod_t$ both tests agree and point at stationarity.

4.3.2 Real activity index

To capture the component of worldwide real activity that stimulates global demand for industrial commodities Kilian (2009) developed the *real activity index*, hence called rea_t. Based on the notion that global economic activity is the most important driver of demand for transport services, the index is constructed based on dry cargo shipping rates. The underlying idea behind this business cycle index is simple: As shown by Stopford (1997) the supply curve of shipping services is relatively flat when demand for freight volumes is low. An increase in global economic activity shifts the demand for freight services. As idle ships are reactivated, the supply curve becomes increasingly steep. At full capacity it becomes effectively vertical and important upward shifts in shipping prices are thus typically associated with aggregate demand pressures in global commodity markets. The opposite effect is true when aggregate demand drops. It should be noted that the level of the index has no intrinsic meaning. Instead, changes in global shipping volumes of raw materials are proportionate to this index Kilian and Zhou (2018, p. 57).

The data series has been retrieved from the homepage of Lutz Kilian and includes the updates addressed in Kilian (2018) based on the methodological critique by Hamilton (2018).[3] For detailed information regarding the construction of the index as well as comparisons with other measures of the global business cycle we refer to Kilian (2009, pp. 1055-1058) and Kilian and Zhou (2018). As can be seen in figure 4.4, a long swing in the time-series can be observed, pointing to a substantial degree of persistence. Persistence in time-series can be an indication for nonstationarity. Therefore, the visual inspection suggests the transformation

[3] https://sites.google.com/site/lkilian2019/home.

Lag order	ERS-test (H_0: unit-root)				KPSS-test (H_0: stationarity)			
	Model with constant		Model with trend		Model with constant		Model with trend	
0	-1.603		-2.382		2.347	***	2.230	***
1	-2.230	**	-3.240	**	1.196	***	1.136	***
2	-1.904	*	-2.804	*	0.813	***	0.772	***
3	-1.895	*	-2.801	*	0.622	**	0.591	***
4	-1.643	*	-2.479		0.507	**	0.482	***
5	-1.741	*	-2.609	*	0.430	*	0.409	***
6	-1.700	*	-2.564		0.375	*	0.357	***
7	-1.438		-2.236		0.334		0.318	***
8	-1.365		-2.146		0.302		0.287	***
9	-1.554		-2.389		0.276		0.263	***
10	-1.602		-2.458		0.255		0.243	***
11	-1.762	*	-2.673	*	0.238		0.226	***
12	-1.783	*	-2.714	*	0.223		0.212	**
Significance:	***=1%,	**=5%,	*=10%					

Table 4.2: Stationarity evaluation of the rea_t series.

Δrea_t as will be used in chapter 7.

The formal test results seem at best mixed regarding the null hypothesis of a unit-root in the series. At a lower lag order the null hypothesis is rejected at medium significance levels between 5% and 10%. The null hypothesis of stationarity is rejected for all lag orders. This can be a result of the fact that the KPSS test is prone to severe size distortions leading to a an overrejection of stationarity. The outcomes of both tests seem surprising as Kilian (2014, p. 457) himself points out that the "real activity index is a business cycle index and stationary by construction".

4.3.3 The real price of oil

The real price of oil is included as the third core variable. It is constructed as the US refiners acquisition price (RAP) of imported crude oil deflated by the US CPI. The oil price series is also retrieved from the EIA's *Monthly Energy Review* table 9.1 while the CPI series is retrieved from the Federal Reserve Economic Database (FRED).[4] Kilian (2009) proposes the RAP of imported crude oil as the best proxy for global oil prices. Blanchard and Gali (2007) note however that different crude oil benchmarks typically result in similar empirical results because of quality-price relationships. The real price of oil series rpo_t is expressed in

[4]https://fred.stlouisfed.org/ with the series code **CPIAUCSL**.

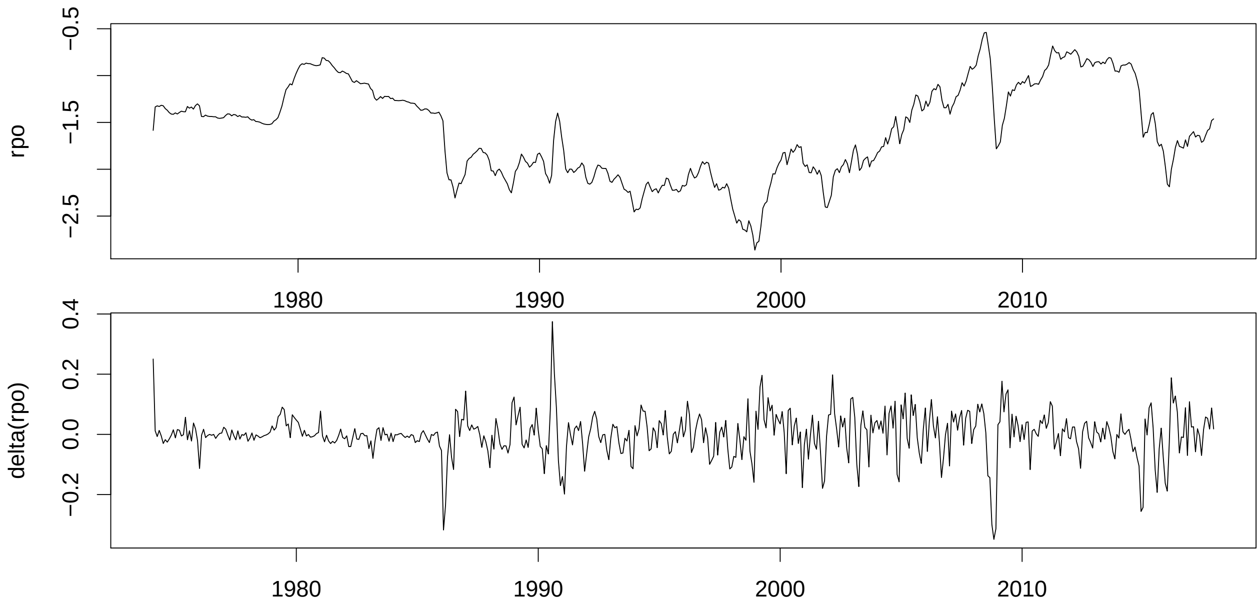

Figure 4.5: rpo_t and Δrpo_t time-series (Source: EIA and FRED).

logarithms.

As can be seen in the upper panel of figure 4.5, the time-series rpo_t shows significant swings, pointing at substantial persistence. This is a sign of unit-root nonstationarity which calls for attention. In chapter 7 we will thus also use the transformation Δrpo_t that can be seen in the lower panel of figure 4.5.

The formal results in table 4.3 mostly reject the null hypothesis of unit-root non stationarity. Keeping in mind, the tendency to overrejection of the KPSS test the stationarity rejections in columns 3 and 4 should not be over-interpreted.

4.3.4 Above-ground oil inventories

Kilian and Murphy (2014) propose to extend the core variable set by a fourth endogenous variable to capture precautionary demand pressures. They follow Hamilton (2009a) in using commercial OECD above-ground crude oil inventory stocks. The above-ground inventories time-series inv_t is based upon those provided by the US EIA which refer to million barrels. Given the lack of data for other countries this measure seems the best proxy for developments of global demand for inventories. Kilian and Lee (2014) use an alternative time-series as compiled and provided by the private company *Energy Intelligence Group* that includes information on other countries, including some important emerging economies such as China.

Lag order	ERS-test (H_0: unit-root)				KPSS-test (H_0: stationarity)			
	Model with constant		Model with trend		Model with constant		Model with trend	
0	-1.676	*	-1.682		8.138	***	7.911	***
1	-2.792	***	-2.800	*	4.091	***	3.977	***
2	-2.528	**	-2.534		2.745	***	2.668	***
3	-2.364	**	-2.371		2.073	***	2.016	***
4	-2.203	**	-2.209		1.671	***	1.624	***
5	-2.144	**	-2.151		1.403	***	1.364	***
6	-1.947	**	-1.952		1.211	***	1.177	***
7	-2.003	**	-2.009		1.067	***	1.038	***
8	-1.952	**	-1.957		0.955	***	0.929	***
9	-1.940	**	-1.945		0.865	***	0.841	***
10	-2.149	**	-2.154		0.791	***	0.770	***
11	-2.214	**	-2.220		0.730	**	0.710	***
12	-2.104	**	-2.110		0.678	**	0.659	***
Significance:	***=1%,	**=5%,	*=10%					

Table 4.3: Stationarity evaluation of the rpo_t series.

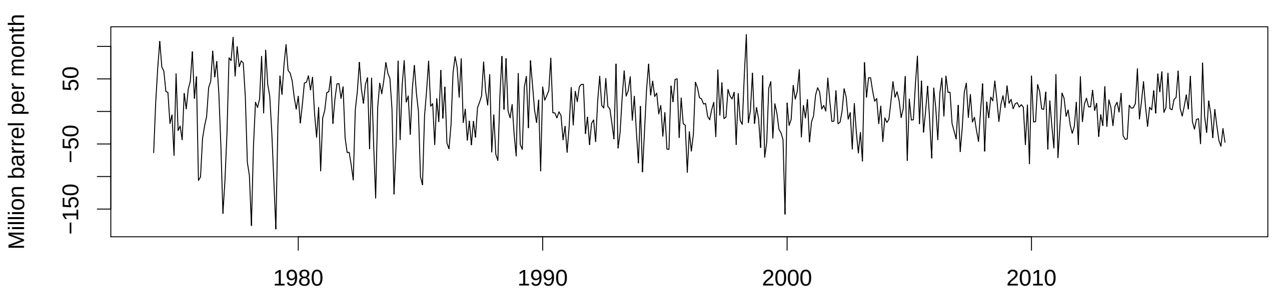

Figure 4.6: inv_t time-series (Source: EIA).

The empirical results remain however very similar.

Due monthly OECD inventory data starting in January 1988, we follow Hamilton (2009a) in constructing the series back to January 1974 by scaling the OECD petroleum stocks over US petroleum stocks growth between January 1974 and December 1987, the latter series also provided by the EIA. Due to a correlation between total OECD and US petroleum inventories of around 80% after 1988, this construction seems plausible. The time-series are expressed in monthly level changes rather than percentage changes to be able to directly compute oil demand elasticity as discussed in the description of the restrictionsfor the Kilian and Murphy (2014) four-variable model.

Lag order	ERS-test (H_0: unit-root)				KPSS-test (H_0: stationarity)	
	Model with constant		Model with trend		Model with constant	Model with trend
0	-8.585	***	-12.349	***	0.203	0.034
1	-5.737	***	-8.531	***	0.157	0.027
2	-4.977	***	-7.616	***	0.135	0.023
3	-5.126	***	-8.138	***	0.126	0.021
4	-4.904	***	-8.084	***	0.128	0.022
5	-3.569	***	-6.002	***	0.140	0.024
6	-3.506	***	-6.036	***	0.155	0.026
7	-2.755	***	-4.796	***	0.181	0.031
8	-2.339	**	-4.115	***	0.214	0.037
9	-1.945	**	-3.435	**	0.250	0.043
10	-1.656	*	-2.932	**	0.283	0.049
11	-1.272		-2.328		0.307	0.054
12	-1.212		-2.230		0.296	0.052
Significance:	***=1%,	**=5%,	*=10%			

Table 4.4: Stationarity evaluation of the inv_t series.

Visually, as can be seen in figure 4.6 and formally, as provided by the test results in table 4.4, no indication of nonstionarity of the the time-series inv_t can be detected.

4.4 Estimation results

In this section we will show and discuss the main results obtained from estimating both structural models presented in section 4.2. We focus on the structural shocks and the responses of the endogenous model variables. Keep in mind that we are interested in the estimation of both models mainly to use the empirical results in the next two chapters.

As shown above, the computation of the structural shocks in both SVAR models requires the estimation of the reduced forms as in equation (4.1) in a first step. We estimate the reduced form models consistently with OLS and subsequently estimate both structural models with ML using the R package **vars**. It provides standard analysis tools in the context of VAR, SVAR and structural vector error correction (SVEC) models. For a full package documentation as well as theoretical overview of the models included we refer to Pfaff (2008). Before presenting the empirical results regarding the structural shocks, we will briefly discuss some diagnostic testing regarding the reduced form errors.

4.4.1 Reduced form residuals diagnostic testing

For both models, we perform standard univariate and multivariate tests for heteroscedasticity, autocorrelation and normality. As the results are very similar for the two reduced form models, we will talk through their results jointly.

To test for heteroscedasticity, we apply the univariate and multivariate autoregressive conditional heteroscedastic (ARCH) tests as proposed by Engle (1982). For all univariate residuals with the exception of $\widehat{u}^{rpo}$ we can strongly reject the null hypothesis of no heteroscedasticity. A visual inspection of the univariate residuals confirms the formal test results. Similarly the null of no heteroscedasticity can be clearly rejected in the case of the multivariate test. Similarly, the univariate and multivariate Jarque-Bera tests for normality based on Jarque and Bera (1980, 1987) and Bera and Jarque (1981) all reject the null hypothesis of normally distributed residuals for both reduced form models.

In the case of autocorrelation we apply univariate and multivariate portemanteau tests based on Ljung and Box (1978). While in the case of the univariate reduced form residual series we cannot reject the null hypothesis of no serial correlation, we strongly reject it in the multivariate case. As has already been discussed in the theoretical part in chapter 3 because of the relationships between the endogenous variables, we indeed expect strong cross-correlation between the residuals. This result is simply verified by looking at the estimated residual variance-covariance matrix $\widehat{\boldsymbol{\Sigma}_u}$ that is not diagonal.

The residual diagnostic testing in case of both reduced form VAR models point to systematic time-varying information contained in the residuals. As in the case of previously applied VAR models in oil research, impulse response analysis based on these residuals can be misleading. A structural modeling of the residuals can thus help overcome these empirical reduced form difficulties by modeling and estimating theoretical relationships. In the following, we will present the estimation results for the structural models as presented in sections 4.2.1 and 4.2.2.

4.4.2 Estimation results for the recursively identified SVAR 3 variable model

Figure 4.7 shows the evolution of the normalized structural shocks between February 1976 and December 2017. For a better overview, the estimated monthly shocks have been aggre-

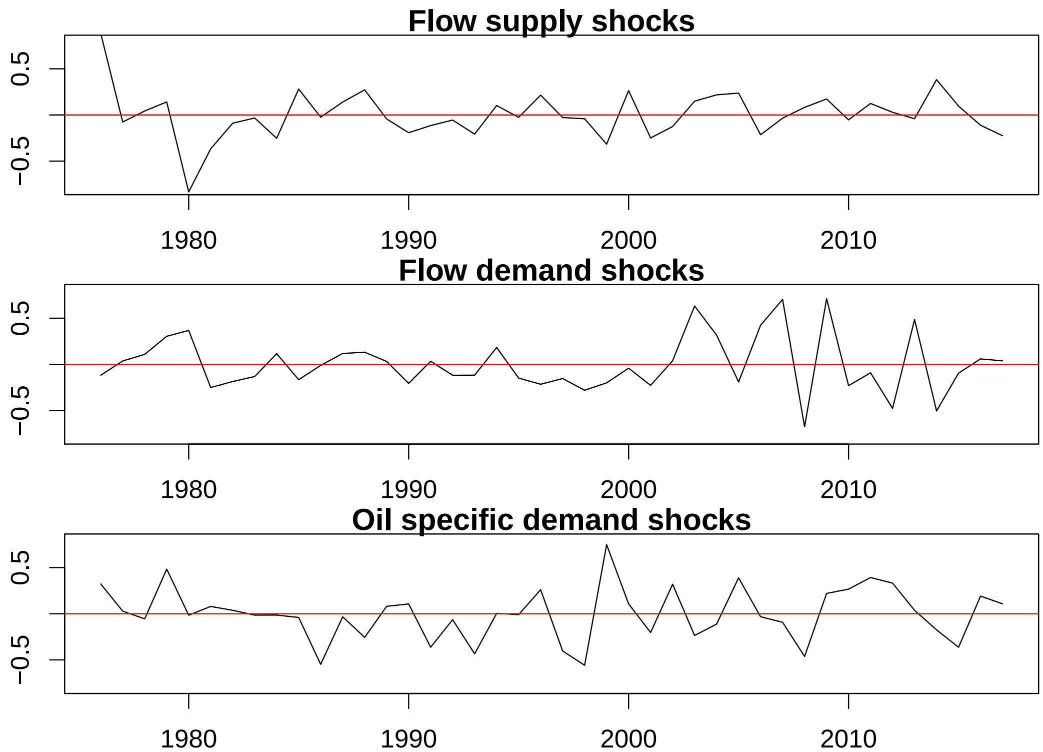

Figure 4.7: Annual averages of structural shocks between 1976 and 2017.

gated to annual averages. Although we use an extended sample up to December 2017 and the updated *real activity index*, we find the same results for the evolution of all three shocks when comparing to the original shock estimates of Kilian (2009, p. 1060).

Some empirical insights can be highlighted. First, global oil markets have been affected by all three shocks simultaneously at any given point in time. Second, negative physical supply disruptions as have been commonly associated with oil price shocks in the literature, appear to be much more attenuated in comparison to flow demand and oil-specific demand shocks. The case of the oil price surges commonly attributed to the Iranian revolution in 1979 is interesting. As can be seen from figure 4.7 the oil price surge experienced by the world in 1979 was preceded by three consecutive years of strong aggregate demand shocks for industrial commodities driving up oil prices. In 1979 we further observe a shock to oil-specific demand, confirming the observations described in chapter 2.2 regarding strong precautionary pressures on oil markets. On the supply side, the 1979 episode is even characterized by a small positive flow supply shock. As higher oil prices induced by the other two shocks gave producers the incentive to supply more crude oil to global consumers.

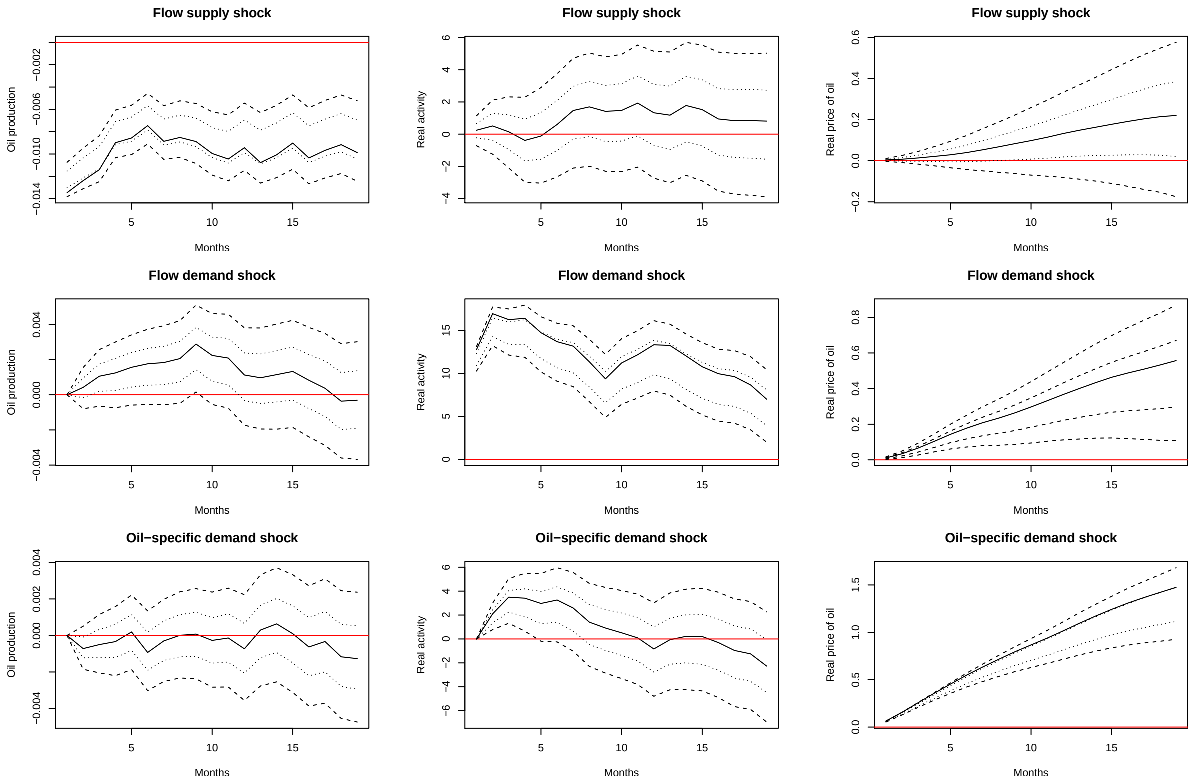

Figure 4.8: Responses to one-standard-deviation structural shocks.

When looking at more recent oil episodes, we recognize that the oil price surge starting in the early 2000s, and peaking in 2007, was mainly due to a combination of aggregate demand shocks for industrial commodities and oil-specific demand shocks. Again, we see that the data doesn't indicate strong unexpected movements on the supply side. The collapse of the oil price in 2009 was mainly due to the great recession and a slump in global economic activity that was accompanied by a decrease in aggregate demand for industrial commodities. The more recent slump in 2014 was due to both, a negative shock to flow demand and a decrease for oil-specific demand. As Kilian (2017) shows, these oil-specific movements are mainly due to the decrease in North American oil inventories that were previously built up following the fracking boom, thus decreasing the US demand for imported crude oil.

Figure 4.8 shows the point estimate responses of global oil production, real economic activity as well as the real price of oil to the estimated standardized structural innovations. The shocks have been normalized such that all induce an oil price increase. The one- and two-standard-error bands have been computed based on a bootstrap with 10.000 runs (Pfaff, 2008, p.22; Efron and Tibshirani, 1993). All the results are consistent with the results in

Kilian (2009).

Upon impact, an unexpected flow supply shock causes a sharp fall in crude oil production. A slight reversion can be observed in the first half year after the shock. This observation is consistent with the notion that supply disruptions in one region are at least partially offset by supply increases in other crude oil producing regions (Kilian, 2009, p. 1062). Within the same period, an oil supply disruption insignificantly affects real economic activity. In fact, as Kilian and Lütkepohl (2017, p. 222) point out, a reasonable assumption in setting the over-identifying restrictions in equation (4.7) would have been $b_{21} = 0$, meaning that the world economy reacts with a delay to disruptions. The estimation results show, that the impact as well as the delayed effects are highly insignificant. This result reinforces arguments that oil price increases due to supply disruptions had weak effects in disrupting economic activity. Finally, a flow supply shock leads to a steady increase in the real price of oil. However, this increase is barely statistically significant within one-standard deviation band.

A flow demand shock because of global aggregate demand leads to a steady increase in oil production. The effect becomes strongly significant after eight months and fades thereafter. Real economic activity responds positively and persistently significant after a flow demand shock. The effect starts to decrease only after a year. The effect on the real price of oil is positive and significant. While small upon impact, much of the effect increases over time and remains highly persistent after 18 months.

When looking at the impulse responses induced by oil-market specific demand shocks, we see that these have no effects on crude oil production. An oil-specific demand shock is however associated with an significant increase in real economic activity in the first five months after the shock. More importantly, an oil-specific demand shock is followed by a steady and statistically significant increase in the real price of oil. In absolute terms this is the most important reaction of the real price of oil.

Figure 4.9 allows us to further evaluate the importance of the three structural shocks by plotting their cumulative effects the real price of oil (in 1980-82 US$) since February 1976 by decomposing the data. A first look at the three panels reveals that historically, both types of demand shocks play a more important role in explaining important oil price shifts than the flow supply shocks. For reading and interpreting the results shown in figure 4.9 it is again helpful to look at different oil episodes mentioned in the historical review. As can be seen, the strong increase in real oil prices in 1979 is mainly due to flow demand and especially

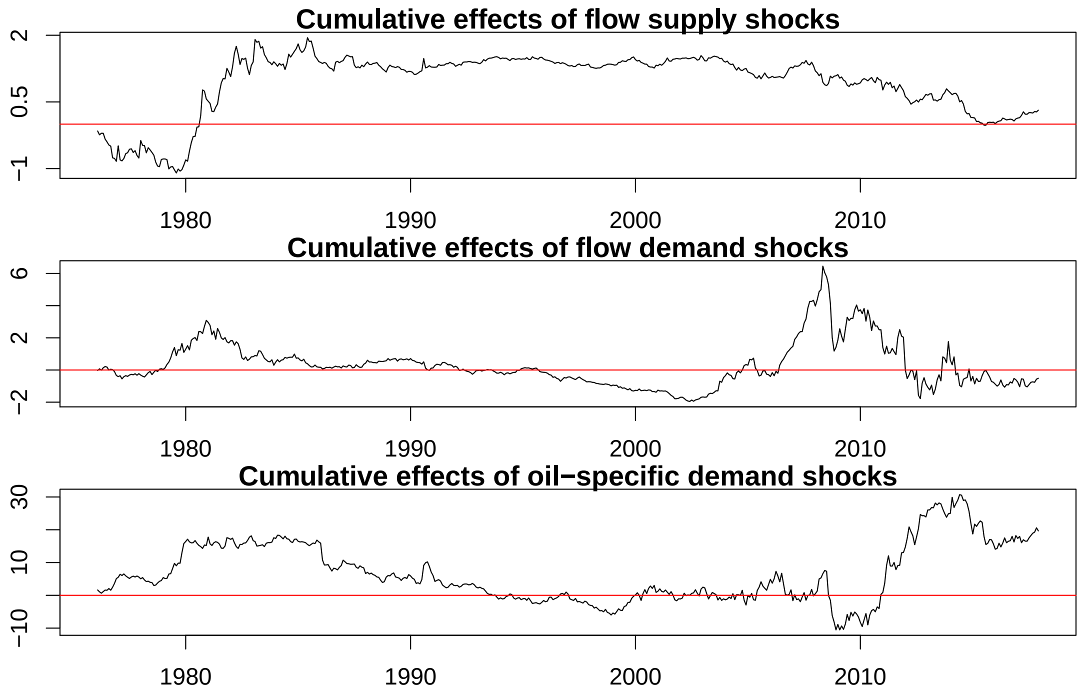

Figure 4.9: Cumulative response of the real price of oil in 1980-82 US$ to the structural shocks in the three-variable model.

oil-specific demand shocks. The cumulative increase from January 1979 to December 1980 amounts to around 10 US$. Flow supply shocks even contributed to a small oil price decrease over the same two years. It is only after the start of the Iraq-Iran War in 1980 that we observe an increase in the real price of oil due to flow supply disruptions. The 1980 oil price increase episode was reinforced by flow demand shocks while oil-specific demand shocks played no further role. The strong increase of the real price of oil prior to 2007 was due to a combination of flow demand shocks due to strong real economic activity and oil-specific demand shocks. When looking at the last episode between 2007 and 2017. We see that the first important drop was due to supply increases, as well as an important drop in flow demand and oil-specific demand. While oil-specific demand recovered, mainly due to a strong recovery of oil markets in emerging economies and especially China, the contribution of flow supply and flow demand shocks was cumulatively negative since 2010. On the one hand, global crude oil production continued to increase, mainly driven by strong output growth in the United States. On the other hand, the global economy was slow to rebound from the Great Recession and Europe was hit by the Euro Crisis.

4.4.3 Estimation results for the sign-restricted SVAR 4 variable model

In order to check the robustness of our results we now present the results for the sign restricted estimation of the four-variable SVAR model based on Kilian and Murphy (2014). In order to reduce the number of iterations necessary to estimate the set of admissible solutions $\tilde{\boldsymbol{B}}^*$ for the structural matrix $\tilde{\boldsymbol{B}}$ we modify the identification loop proposed by Kilian and Murphy (2012, 2014) and introduced in section 4.2.2 on page 40 by expanding the search algorithm to make use of the *svar* function for R as follows:

1. Draw an 4×4 matrix $\boldsymbol{K}$ of NID(0,1) random variables. Obtain the $\boldsymbol{QR}$ decomposition of $\boldsymbol{K}$ so that $\boldsymbol{K} = \boldsymbol{QR}$ with $\boldsymbol{QQ'} = \boldsymbol{I}_4$ and $\boldsymbol{R}$ being upper diagonal.
2. Define $\boldsymbol{D} = \boldsymbol{Q}'$ and construct the structural impact response matrix $\tilde{\boldsymbol{B}} = \boldsymbol{BD}$ to compute the implied structural impulse responses.
3. Before checking if all conditions are satisfied by $\tilde{\boldsymbol{B}}$, remove six random entries and optimize $\tilde{\boldsymbol{B}}$ by ML as in Pfaff (2008, p. 4).
4. If all restrictions (sign restrictions on $\tilde{\boldsymbol{B}}$, elasticities and impulse response conditions) are satisfied, retain $\tilde{\boldsymbol{B}}$ else discard $\tilde{\boldsymbol{B}}$.

After 3.000.000 repetitions, we obtain a solution set with 46 admissible structural models. In figure 4.10, we show the impulse response functions corresponding to the whole solution set. As expected, it becomes apparent that verifying sign-restricted conditions in addition to conditions that are market-specific, yields models with significant differences in their implied impulse response functions. We point out that in empirical applications, researchers often treat the vector of pointwise medians as if it were a pointwise estimate. As Fry and Pagan (2005, 2011) indicate, this practice is misleading as there is no reason for one model resulting in the median responses for all impulse response functions simultaneously. We thus choose the benchmark model with a short-run price elasticity of oil demand closest to the value of -0.26 as postulated by the literature (Kilian and Lee, 2014; Kilian and Murphy, 2014; Kilian, 2017).

Figure 4.11 shows the point estimate responses of global oil production, real economic activity, the real price of oil as well as global crude oil inventories to the estimated standardized structural innovations of our benchmark estimate. Again, the shocks have been normalized such that all induce an oil price increase. Pointwise 68% and 95% posterior error bands may be constructed, following a modified procedure as proposed by Uhlig (2005). However, as Kilian and Murphy (p. 1183 2012) point out, these posterior error bands do not correspond

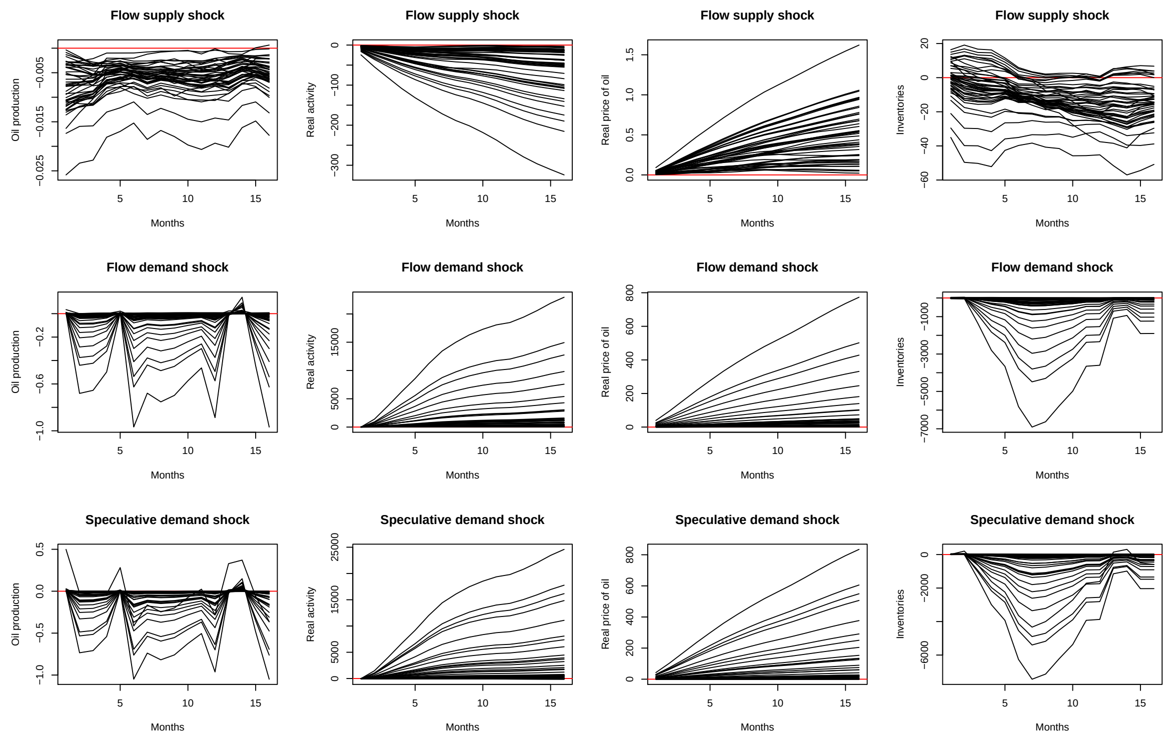

Figure 4.10: Impulse responses corresponding to the whole solution set.

to classical error bands, not even asymptotically. Thus, we focus on the pointwise impulse responses of the chosen benchmark model to verify its plausibility in comparison to the three-variable model as well as to results in Kilian and Murphy (2014).

In addition to the quantification of the effects on oil inventories, the following differences are identified in comparison to the impulse responses as implied by the three-variable recursively identified model (see figure 4.8). As the interpretation of the idiosyncratic oil-specific demand shock is hard to carry out in this model, we focus on the first three shocks. Given a flow supply shock, real economic activity reacts negatively. Furthermore we identify a direct negative effect on oil inventories as these are used to cushion against the global production disruptions. The effect of a flow demand shock resulting from strong global activity on crude oil production is almost twice as strong in this second model. Again, we see that crude oil inventories are used to satisfy a part of the increased crude oil demand resulting from a flow demand shock. As crude oil production increases, the change in inventories starts to reverse after four months. Finally, a speculative demand shock is followed by a reduction in crude oil production and an increase in the real price. This is in line with the argument, that in anticipation of future flow supply disruptions or price increases due to other conditions,

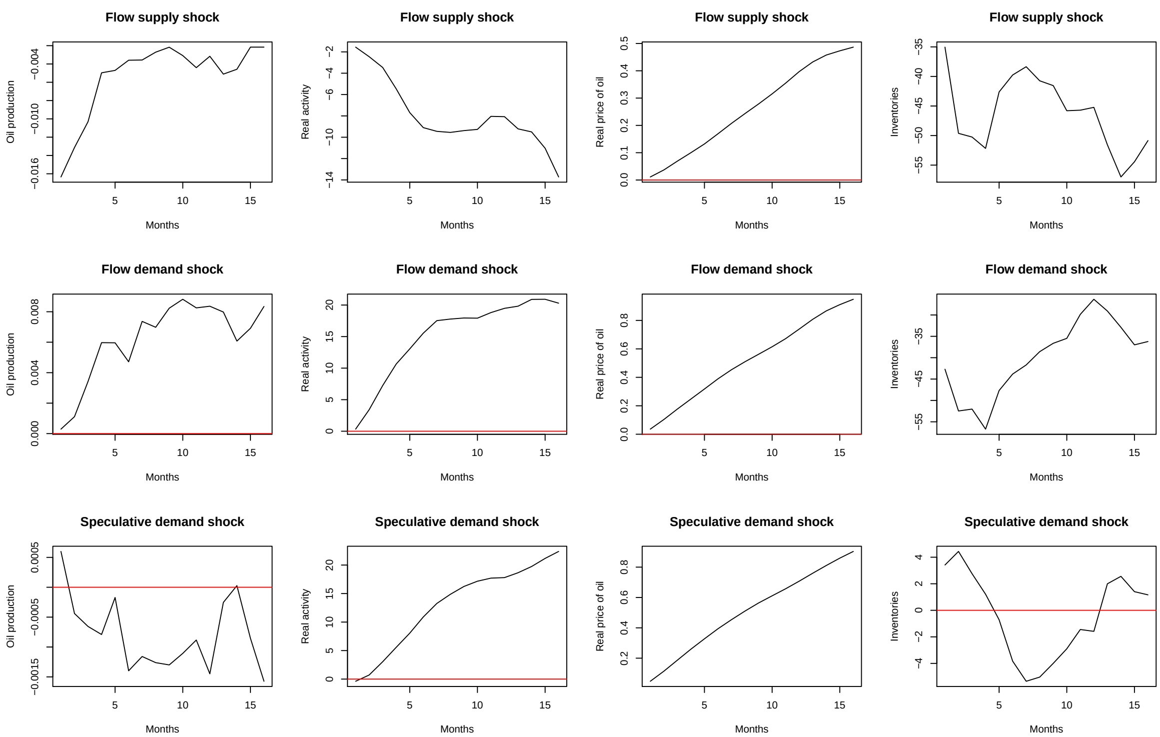

Figure 4.11: Impulse responses to one-standard-deviation structural shocks in the benchmark model.

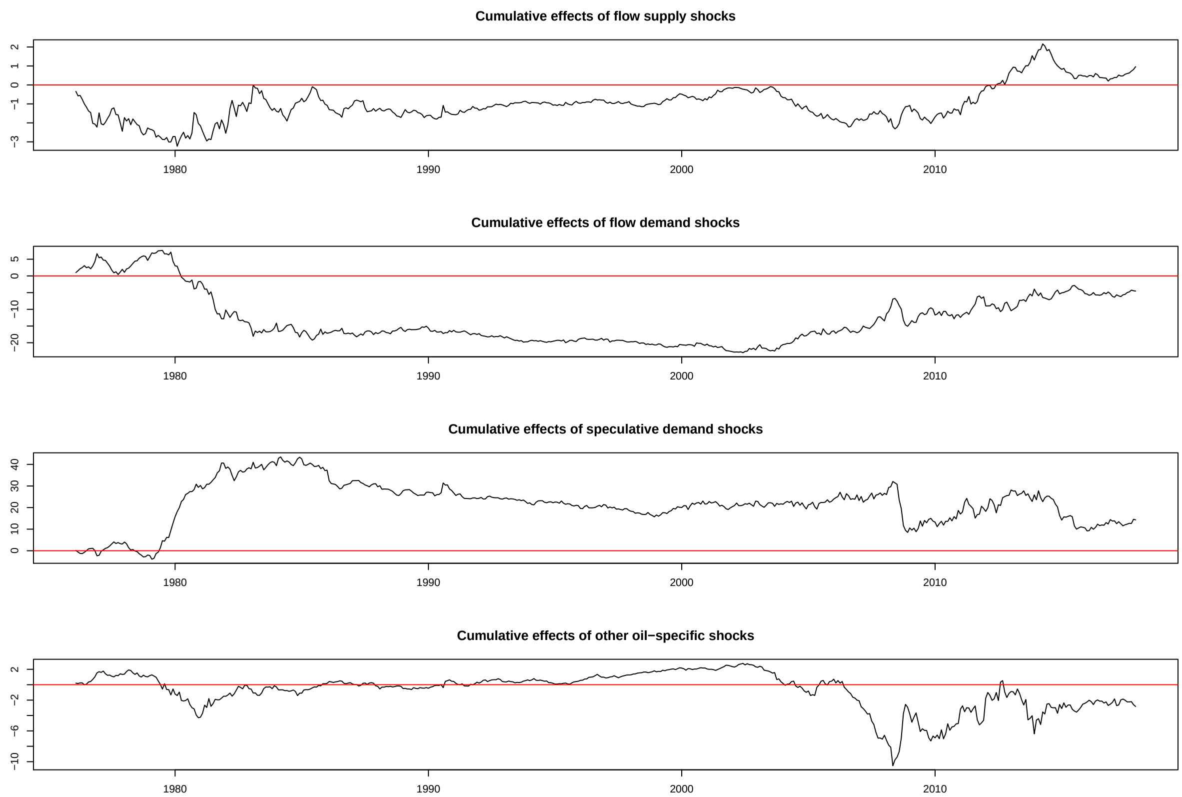

Figure 4.12: Cumulative response of the real price of oil (in 1980-82 US$) to the structural shocks in the four-variable model.

this type of demand will increase. We thus first observe an positive change in inventories within the first five months, followed by a decrease to compensate for the reduced production.

As can be seen from figure 4.12, the historical decomposition of the cumulative shock effects on the real price of oil offers similar insights as in the case of the three-variable model. First and foremost, as in the case of the three-variable model, precautionary demand shocks in anticipation of future disruptions or other crude oil price increases play the most important role in explaining important and unexpected absolute shifts in the real price of oil. While the effects of flow supply shocks remain qualitatively small, flow demand shocks due to strong global demand for industrial commodities remain important in explaining shifts in the real price of oil. We should however point out the main difference in comparison to the results in figure 4.9. While model 1 attributed a negative cumulative effect on the real price of oil to flow supply shocks since the Great Recession, we see the inverse development in the case of the second model. Also, global real activity seems to have recovered much stronger in the second model after 2010, driving prices upwards.

We suspect that the underestimation of the more recent inventories data, mainly due to missing data regarding China as discussed in Kilian and Lee (2014, pp. 74-76), underestimates the effects of speculative demand shocks and overestimates the effects of all other shocks in the post Great Recession period. The data misspecification didn't seem to matter when OECD countries were the main flow demand and precautionary demand drivers in global oil markets. However, it seems to matter when global demand for oil rises due to Chinese flow demand or Chinese speculative demand for example. This result should be kept in mind for future applications of this model. The question whether to model precautionary demand should depend on the accuracy of available inventories data.

4.5 Summary

In this chapter we estimated two global SVAR models of oil. In both cases the goal is to disentangle the underlying causes of unexpected oil price movements by defining different structural shocks. First, we estimated the recursively identified three-variable SVAR model proposed by Kilian (2009) that includes flow supply shocks, flow demand shocks and oil-specific demand shocks. Second, we estimated the sign-restricted four-variable SVAR model proposed by Kilian and Murphy (2014) that further splits oil-specific demand into speculative demand shocks and other oil-specific demand shocks.

Our estimations based on a larger data sample confirm the main results of Kilian (2009), Kilian and Murphy (2012) and Kilian and Murphy (2014):

- Contrary to early analyses as seen in chapter 3, flow supply shocks in the form of unexpected disruptions to the physical availability of oil, played a lesser role in explaining unexpected oil price movements.
- Flow demand shocks because of the global business cycle and oil-specific demand shocks are at the root of most unexpected oil price movements.
- As confirmed by the four-variable SVAR model estimates, oil-specific demand is primarily due to speculative demand. Market participants anticipate future price increases and consequently build up inventories for future consumption resulting in higher prices.
- The estimation of the four-variable SVAR model poses some challenges that should be further investigated. We observed considerable differences in the solutions chosen by the drawing algorithm. Furthermore the choice of OECD inventories as a proxy for

global oil inventories might bias the estimation results at the end of the sample as it excludes important countries such as China.

In the next two chapters, we will apply the results presented in this chapter to three empirical analyses. First in chapter 5, in the spirit of Kilian (2009), we will estimate and compare whether and how US and German macroeconomic aggregates react to the three different oil price shocks resulting from the three-variable SVAR model. As has been shown in the literature review, most empirical studies have focused on the United States, while additionally lacking the disentanglement of the underlying causes of oil price shocks, as presented in this chapter. For the evolution of macroeconomic aggregates it might matter if a given oil price increase results from supply disruptions or an increase in global aggregate demand. The second empirical analysis in chapter 6 follows and complements the spirit of Kilian (2017) and studies the effects of the economic and political sanctions imposed on the Iranian oil and energy sector that we presume as exogenous. We will evaluate the forecast performance of the different SVAR models presented in this chapter given the construction and use of an exogenous Iran supply shock series. We will then construct different counter-factual crude oil production series for Iran in order to forecast the effects of the two considered sanction rounds on the global price of oil. Important observations with regard to expectations will be made. Finally, we will evaluate the forecast performance of the base reduced form VAR model against models with additional variables selected by different sparse selection methods in chapter 7.

Chapter 5

Revisiting the oil price-macroeconomy relationship: A comparison between Germany and the United States

As already mentioned in chapter 3, early research focused on the question if and how oil price changes and important oil price shocks affect macroeconomic aggregates. We follow Kilian (2009) in using the three orthogonal structural shocks resulting from the recursively identified three-variable SVAR model in order to disentangle the underlying causes of an oil price surge and their potentially different effects on macroeconomic aggregates. Our analysis complements the literature in following in the ways: First, the few studies that include Germany with regard to the effects of oil price changes on macroeconomic aggregates do not disentangle the underlying causes of the price changes. We thus expect that the German economy, characterized by a heavy dependence on imported oil as well as strong trade linkages to the rest of the world, reacts differently if the cause of an oil price increase is due to a positive global aggregate demand shock for industrial commodities or a negative oil supply shock. Second, in addition to quarterly GDP data, we also investigate the effects of monthly time series for both the US and Germany. In comparison, Kilian (2009) uses his SVAR framework to analyze the impact of the three different oil price shocks on quarterly US GDP growth and CPI inflation by aggregating the monthly structural shocks into a quarterly format.

In the following, we will first present the empirical framework as well as the data. We will then show and describe the empirical results before finishing the analysis by comparing the differences between both countries.

5.1 Empirical framework and data

As Kilian (2009, p. 1056) notes, once we estimate the structural shocks from the three-variable model, we can use them to determine the reaction of macroeconomic aggregates to oil price shocks. Let's recall from section 4.2.1 that all three structural shocks result in an oil price increase. The first shock estimate was defined as a flow supply shock, $\widehat{\varepsilon}_t^{\text{flow supply shock}}$ standing for unexpected changes to the physical availability of crude oil. The second shock estimate is referred to as a flow demand shock $\widehat{\varepsilon}_t^{\text{flow demand shock}}$ due to global demand for industrial commodities (including crude oil) resulting from the global business cycle. Finally, $\widehat{\varepsilon}_t^{\text{oil-specific demand shock}}$ refers to price changes that are not accounted for by the first two structural shocks such as precautionary demand or technological shocks. Thus, y_t be the macroeconomic variable of interest and $\widehat{\varepsilon}_{jt}$, $j = 1, 2, 3$, the structural shock estimates. Then we are able to determine the effects of each shock on y_t by estimating following regressions:

$$y_t = \beta_{j0} + \Sigma_{0=1}^{h} \Phi_{ji} \widehat{\varepsilon}_{jt-i} + u_{jt} \tag{5.1}$$

where u_{jt} are the error terms. In this model Φ_{jh} consequently corresponds to the impulse response coefficients at horizon h. Given the possibility of serial correlation regarding the residuals, we follow Kilian (2009) by using block bootstrap methods in order to estimate the confidence intervals. We thus construct the one- and two-standard deviations error bands using a block size of 4 and 10.000 replications for each regression. Another potential issue is whether the three structural shocks are predetermined with respect to US and German macroeconomic aggregates. Kilian (2009, pp. 1065-66) argues that the shocks are predetermined when using US quarterly data, i.e. they do not react to US macroeconomic shocks within the quarter. We are thus confident that this will also apply to monthly US data. As the German economy is much smaller and less oil intensive than the US economy, the assumption is easily justified with regard to German macroeconomic aggregates on monthly and quarterly frequency. The \$/€ spot exchange rate might be an exception as markets generally are fast to react to new information. As the exchange rate series is of no central importance to the underlying analysis, this should be kept in mind primarily when comparing German series originally expressed in € and transformed to \$ for comparison purposes.

Given that GDP is only available at a quarterly frequency, we aggregate the monthly shock estimates to quarterly shocks by constructing the mean in each quarter. The same regressions as in (5.1) are then estimated with quarterly GDP data. The maximum impulse horizon is set to three years in accordance withKilian (2009), thus $h = 36$ for the monthly and $h = 12$ for the quarterly regressions.

Variable	Frequency	FRED Series code	Original Source
\$/€ exchange rate	m	CCUSSP01DEQ650N	OECD
German CPI	m	DEUCPIALLMINMEI	OECD
Real German exports (€)	m	XTEXVA01DEM664S	OECD
Real German imports (€)	m	XTIMVA01DEM664S	OECD
Real German exports (\$)	m	XTEXVA01DEM664S	OECD
Real German imports (\$)	m	XTIMVA01DEM664S	OECD
German industrial production index	m	DEUPROINDMISMEI	OECD
German unempoyment rate	m	LMUNRRTTDEM156S	OECD
US CPI	m	CPIAUCSL	US BLS
US exports	m	XTEXVA01USM667S	OECD
US imports	m	XTIMVA01USM667S	OECD
US industrial production index	m	INDPRO	US BLS
US unemployment rate	m	UNRATE	US BLS
German CPI	q	DEUCPIALLQINMEI	OECD
Real German GDP	q	DEUGDPNQDSMEI	OECD
Real US GDP	q	GDPC1	US BEA

Table 5.1: Used time series and sources.

Besides the sign and significance discussion of the cumulative impulse response functions as estimated by equation 5.1, we complement the significance evaluation by using a Wald test. The following null hypothesis is thus tested for every German and US macroeconomic series regarding all three shocks:

$$H_0 : \Sigma_{i=0}^{k} \Phi_{ji} = 0, \tag{5.2}$$

where k stands for the horizon up to which the cumulated effect should be tested. For monthly series k is set to $k = 12, 24, 36$ and for quarterly GDP series k is set to $k = 4, 8, 12$.

The data sources for the macroeconomic time series used in the analysis is shown in table 5.1. The sample period for the monthly series is February 1976 to December 2017 and for the quarterly series Q1 1976 to Q4 2017. We note, that all nominal series are transformed to real values using the corresponding national CPI. The German series expressed in \$ are calculated using the \$/€ exchange rate. All series are seasonally adjusted and expressed in growth rates in order to avoid non-stationarity issues. We test the unit-root null hypothesis using the ERS test of Elliott et al. (1996) and we test the reverse null hypothesis of stationarity using the KPSS test of Kwiatkowski et al. (1992). For all series, the test results point to stationarity.

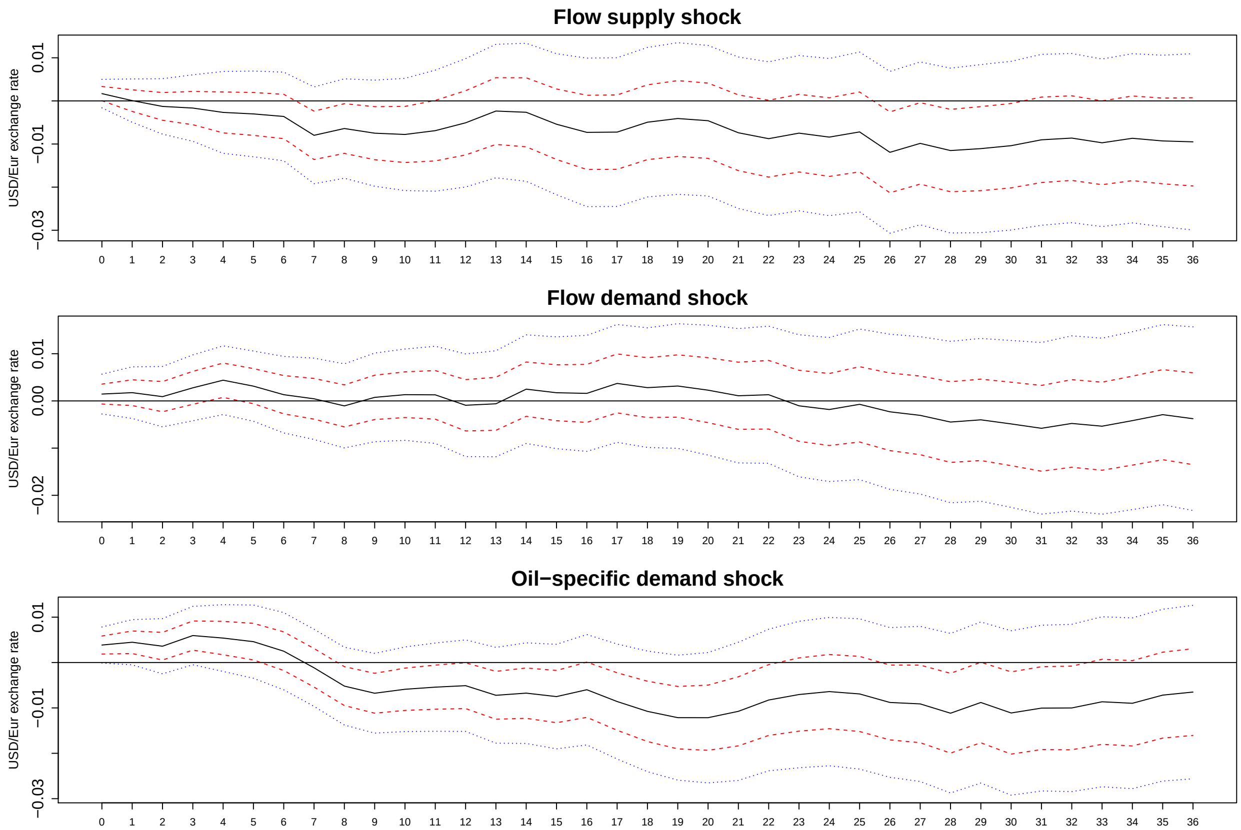

Figure 5.1: Cumulated responses of the \$/€ exchange rate changes to each structural shock. (Point estimates with one- and two-standard error bands)

5.2 Empirical results

Before looking at the specific results we first indicate that for the graphical analysis we evaluate the statistical significance based on the one and two standard error bands in order to follow the nomenclature in the literature (Kilian, 2013; Kilian and Murphy, 2012, 2014). Under the assumption of normally distributed parameter estimates, these error bands roughly correspond to 68% and 95% significance levels. We will mostly concern ourselves with results within the two standard deviations and judge these as significant (in some rare cases we talk about weak significance when results are only within the one standard error bands). We first start with the empirical results regarding the monthly macroeconomic series. At first we look at the reaction of the \$/€ exchange rate following a crude oil supply shock, an aggregate demand shock for industrial commodities and an oil-specific demand shock. The results might help to understand the differences in impulse-responses given the expression of real values in € or US\$.

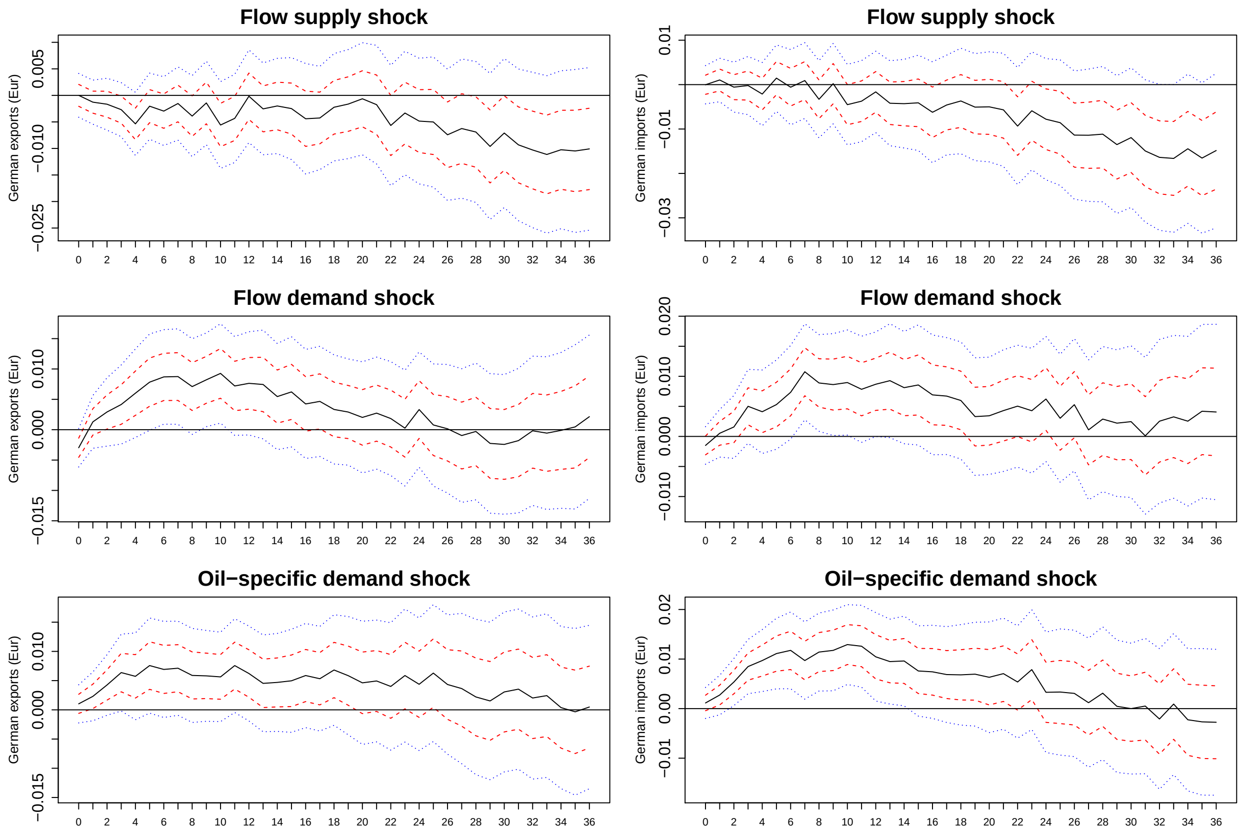

Figure 5.2: Cumulated responses of German exports (left) and imports (right) changes (in Euro) to each structural shock. (Point estimates with one- and two-standard error bands)

From figure 5.1 we can see that the $/€ exchange rate only very weakly reacts to an oil supply shock in the sense that the U$ appreciates vis-à-vis the € after six months. This weak effect seems to dissipate 12 months after the shock but remains for the most part weakly significant. With regard to a flow demand shock, we see no significant effect on the exchange rate. Given an oil-specific demand shock however, we observe again a weak statistically significant appreciation of the $ vis-à-vis the € starting 8 months after the shock. The effect remains mostly weakly significant until 21 months after the shock. This is the first interesting result. When we observe oil price changes, only an oil-specific demand shock will result in an appreciation of the $ compared to the € again highlighting the importance of precautionary demand in oil markets. Higher demand for crude oil specifically (in comparison to aggregate demand for industrial commodities that might be traded in other currencies as well) results in higher crude oil prices, thus increasing the global demand for $.

In Figure 5.2 we show the cumulative responses of the growth rate of German exports and imports expressed in Euro. Following an oil supply shock, we see no significant effect on

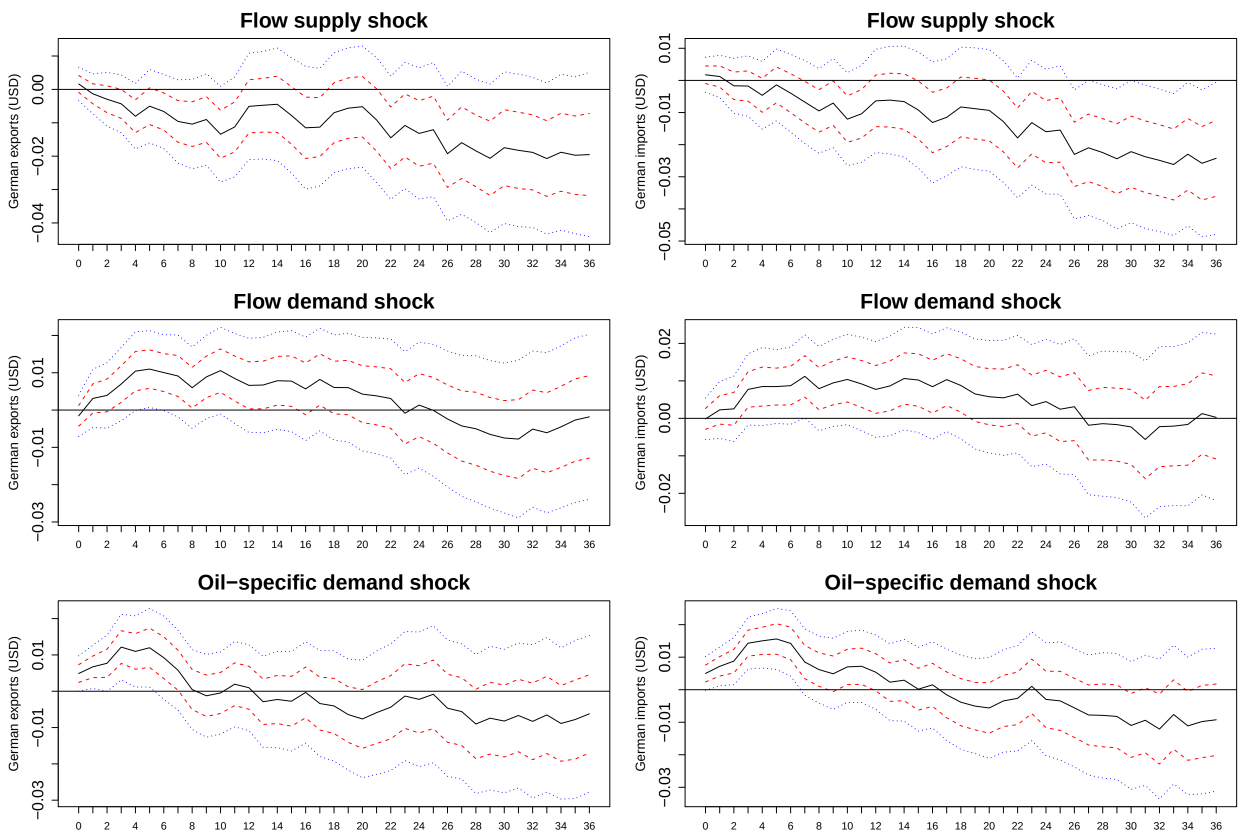

Figure 5.3: Cumulated responses of German export (left) and import (right) changes (in USD) to each structural shock. (Point estimates with one- and two-standard error bands)

the growth rate of exports and imports. When facing a flow demand shock we observe an immediate and significantly positive effect on the growth rate of exports up to 14 months after the shock. To a lesser degree we observe a positive and significant effect on import growth between 5 and 15 months. Both observations point to the important role of trade for the German economy that is heavily export oriented. Moreover, we see a similar reaction when looking at oil-specific demand shocks. Although statistically weaker, German exports react positively for the most part of the first 24 months. Imports on the other hand see a significant increase in their growth rate between 2 and 10 months after the shock indicating that resulting higher oil-prices directly affect the trade balance in a negative way.

This picture has to be complemented by the cumulative impulse response reactions of exports and imports expressed in US$ as seen in figure 5.3. Two main differences become apparent. First, a flow supply shock results in a significant negative effect on the growth rate of imports after around 26 months. Second, the effects of a flow demand shock on export and import growth rates, although remaining significant and positive, are of much shorter duration. The same positive and immediate effect on the import growth rate can be observed following an oil-specific demand shock. It is thus important, to take into account the currency when analyzing the behavior of international macroeconomic variables in the context of oil price shocks.

Figure 5.4 shows the cumulative responses of German industrial production and unemployment growth rates to all three shocks. The responses seem complementary and consistent with each other. Given an oil supply shock, we observe a negative and significant effect on changes in industrial production after 25 months. On the other hand, oil supply shocks immediately induce higher and significant unemployment rates over the whole impulse response horizon. A flow demand shock is followed by a brief positive and significant impulse to industrial production between 7 and 10 months after the shock. The same can be said regarding the unemployment rate following a flow demand shock. We identify a decreasing and significant effect between 7 and 10 months after the shock. Finally, industrial production reacts significantly and positive to an oil-specific demand shock in the few eight months, the effect weakens and becomes insignificant after 21 months. While we observe only a brief, weakly significant and negative effect on unemployment between 7 and 9 months after the shock.

Finally, when looking at the evolution of German CPI inflation following the three oil price shocks in figure 5.5, two main observations can be made. First and foremost, CPI inflation starts to become and remains significantly higher around 6 months after an oil supply shock.

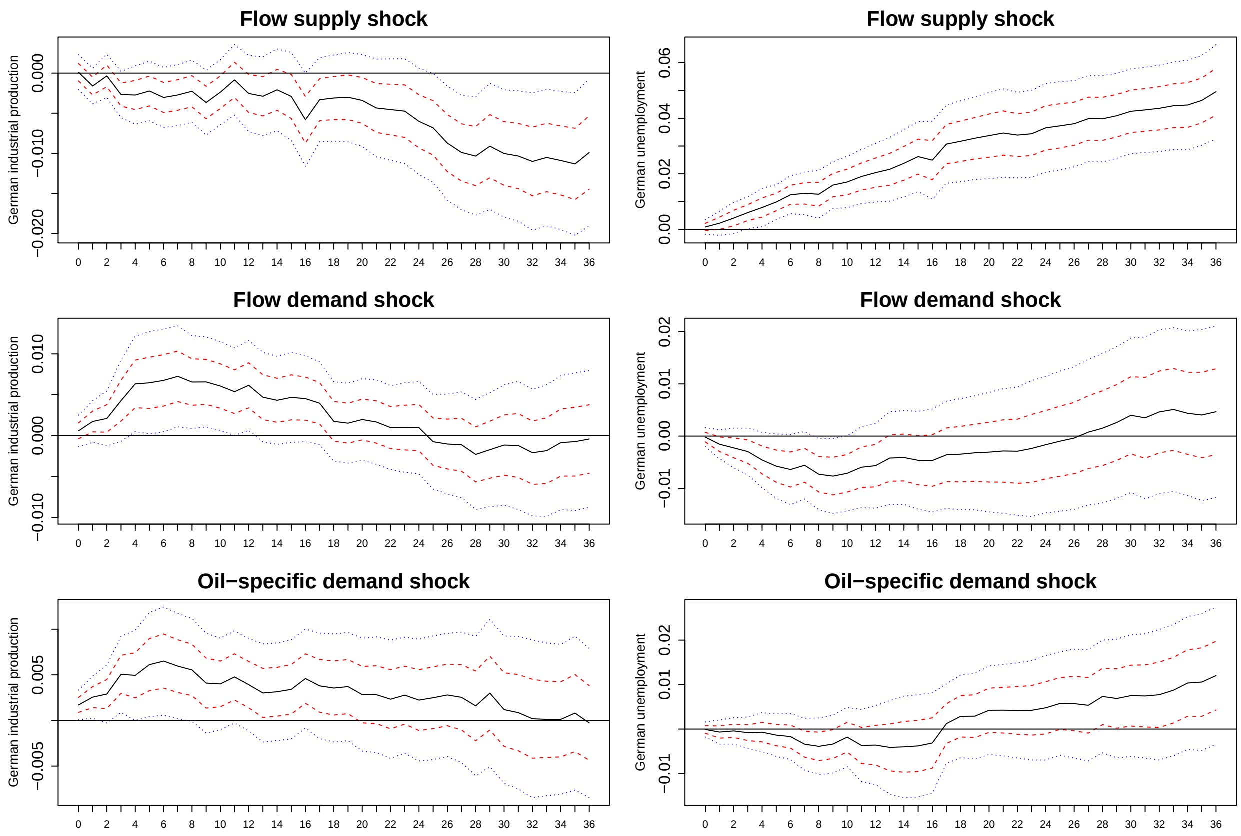

Figure 5.4: Cumulated responses of German industrial production (left) and unemployment (right) changes to each structural shock. (Point estimates with one- and two-standard error bands)

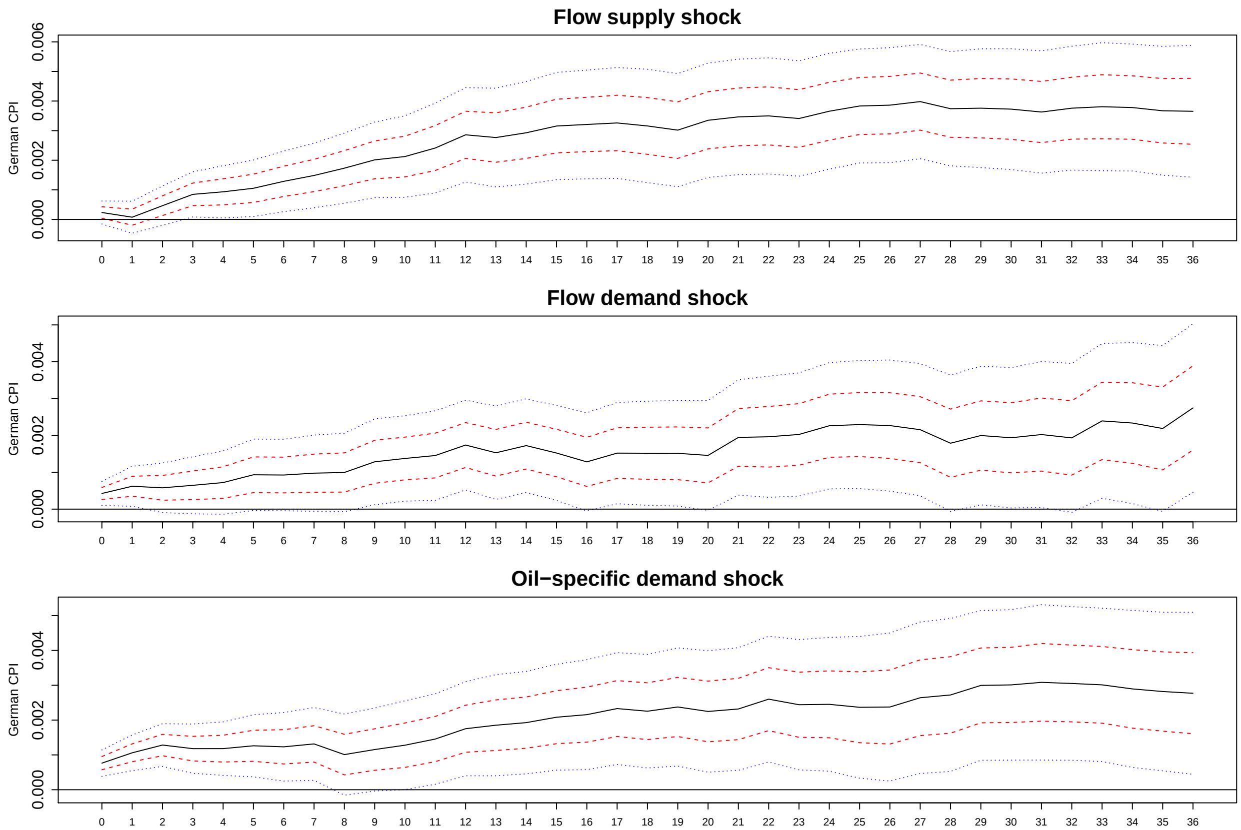

Figure 5.5: Cumulated responses of German CPI inflation to each structural shock. (Point estimates with one- and two-standard error bands)

Second, the reaction of CPI inflation is similar when flow demand or oil-specific demand shocks are responsible for the oil price increases. We identify immediate and statistically significant upward price pressures after both types of shocks. These effects remain significant for the most part over the evaluation horizon of 36 months.

The response of US trade to the three structural shocks can be seen in figure 5.6. As can be seen in the first row, US exports react negatively and statistically significant after 23 months following an oil supply shock. Imports on the other side show no significant reaction thereafter implying that the trade balance as a whole seemingly deteriorates. Regarding oil price increases resulting from a flow demand shock, we observe a strong and statistically significant positive effect on exports in the first 23 months. While shorter, a positive and significant effect is identified on import growth between 2 and 12 months after the shock. Regarding oil-specific demand shocks, we first observe an increase in exports that is statistically significant between 4 and 8 months after the shock. The effect reverses and becomes negative and significant at the end of the impulse response horizon. Imports on the other hand see a

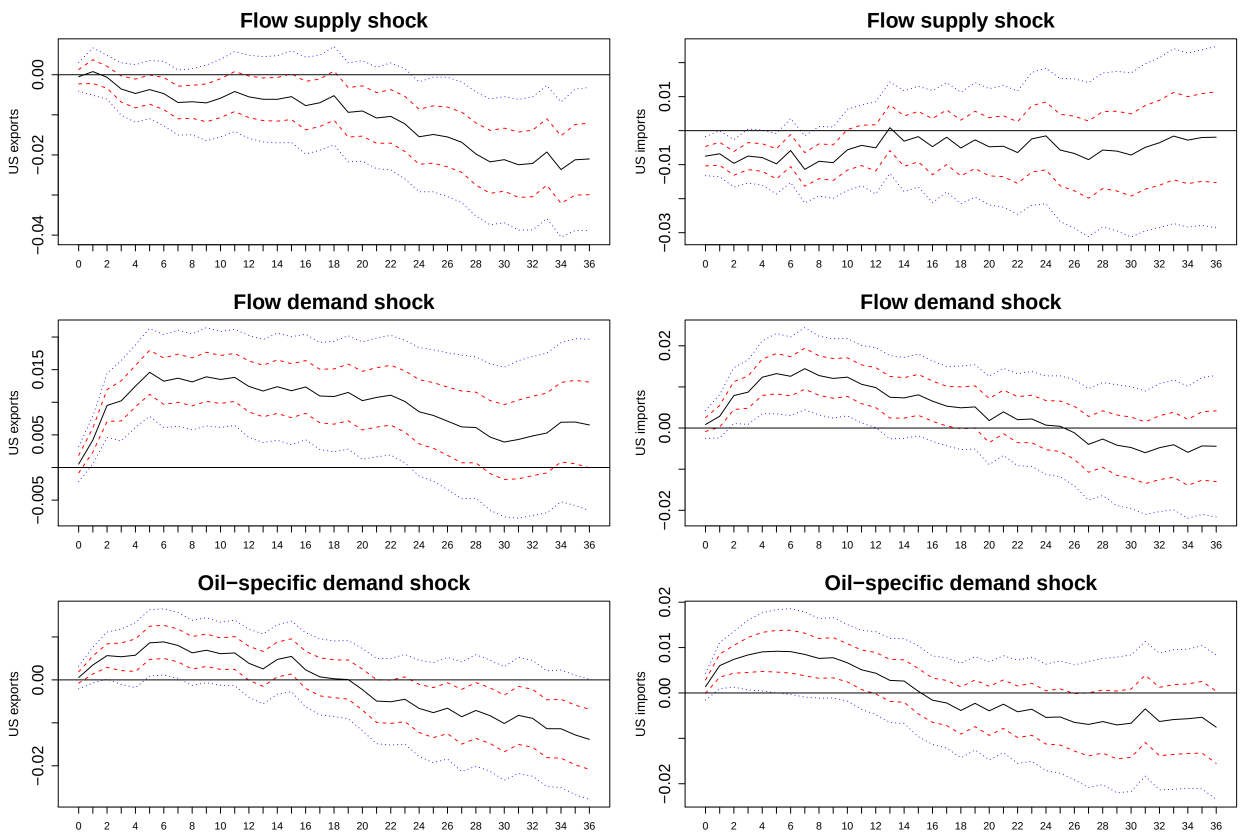

Figure 5.6: Cumulated responses of US export (left) and import (right) changes to each structural shock. (Point estimates with one- and two-standard error bands)

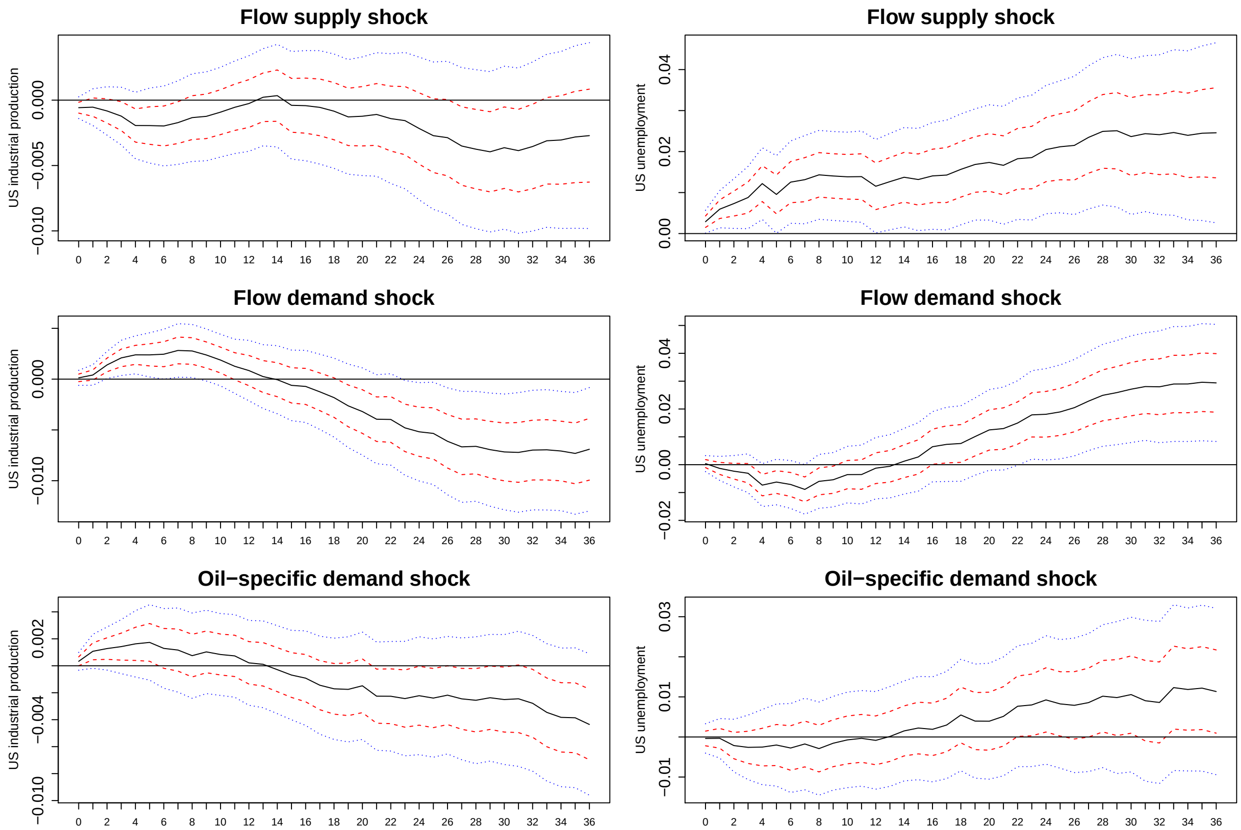

Figure 5.7: Cumulated responses of US industrial production (left) and unemployment (right) changes to each structural shock. (Point estimates with one- and two-standard error bands)

brief increased growth rate in the first three months following an oil-specific demand shock. Afterwards the effect becomes insignificant.

The response of US industrial production and unemployment growth rates can be seen in figure 5.7. As in the case of Germany, both variables react in a complementary way. Following an oil supply shock, we see no significant effect on industrial production. Unemployment growth on the other hand, is significantly higher 18 months after the shock occurs. It remains statistically significant until the end of the impulse response horizon. Following an oil price shock that is due to higher global demand for industrial commodities, we first see a statistical significant increase in industrial production growth and a statistically significant decrease in the growth rate of unemployment in the first 8 and 7 months respectively. For both variables the effect reverses and becomes and remains statistically significant after 22 months following the shock. In the case of oil-specific demand shocks, we observe no significant effect on both variables.

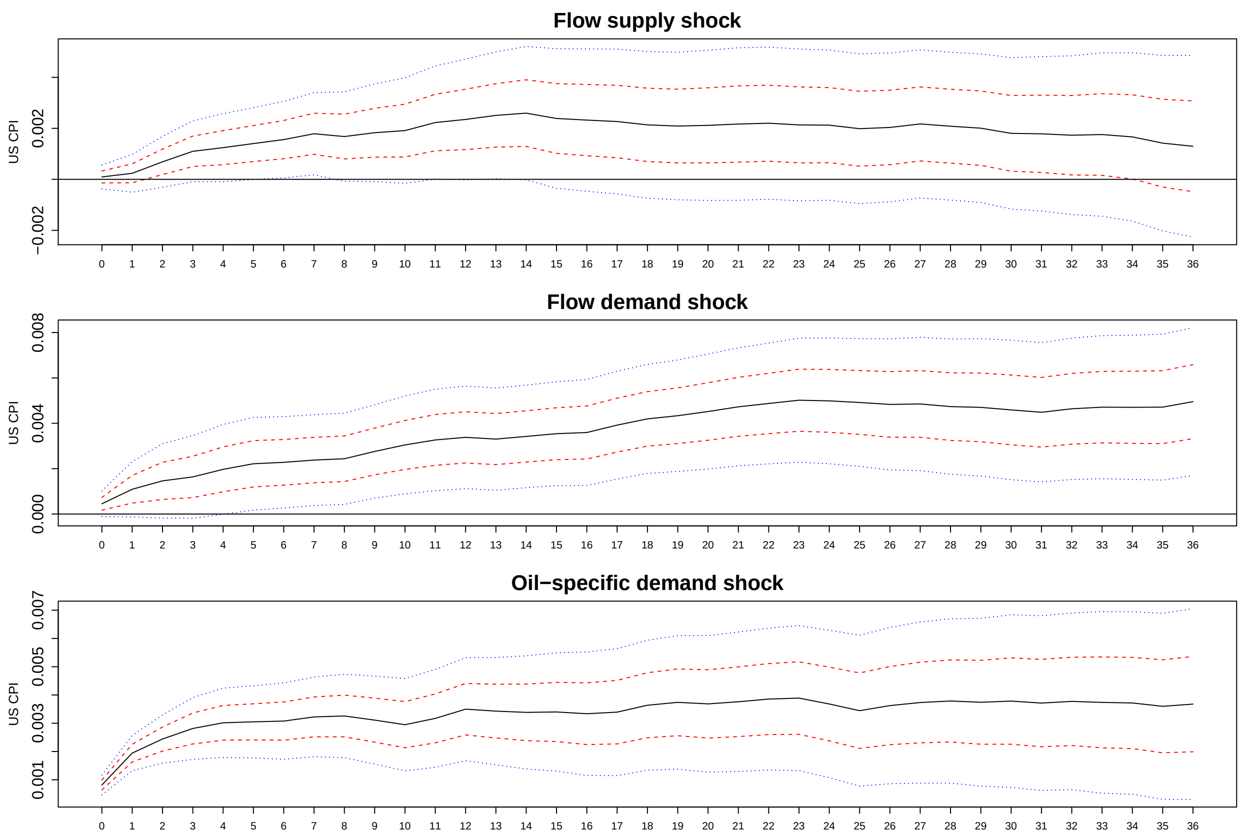

Figure 5.8: Cumulated responses of US CPI inflation to each structural shock. (Point estimates with one- and two-standard error bands)

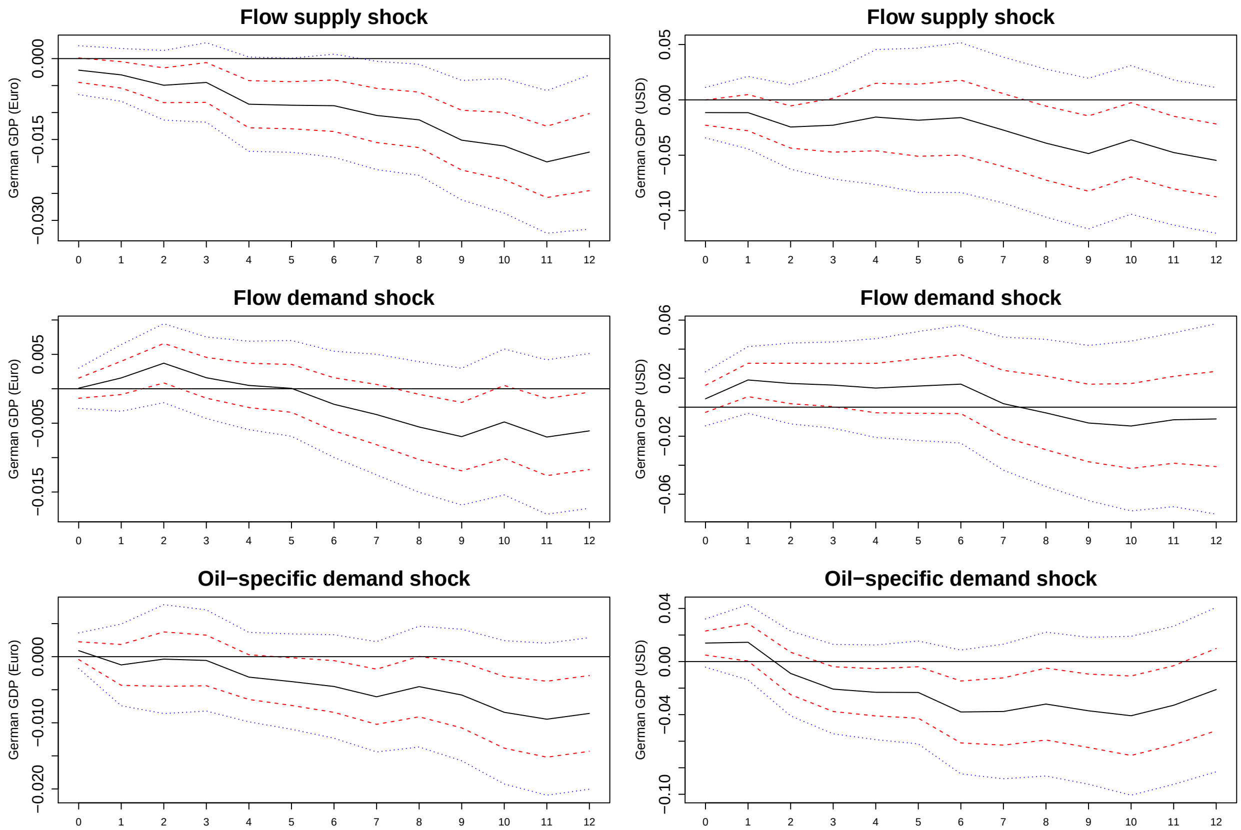

Figure 5.9: Cumulated responses of German GDP growth (left in Euro, right in USD) to each structural shock. (Point estimates with one- and two-standard error bands)

The reaction of US CPI inflation to the structural oil price shocks can be seen in figure 5.8. Following an oil supply shock, we see a positive and statistically significant effect on CPI inflation within the first 14 months. The effect becomes and remains insignificant thereafter. In the case of a flow demand shock, CPI inflation becomes and remains significantly higher 4 months after the shock. The effect of oil-specific demand shocks is also positive and immediately statistically significant. It fades and becomes insignificant 30 months after the shock.

Looking at the results from regressions using quarterly data, we first turn to the the cumulated response of German GDP growth over 12 quarters as shown in figure 5.9. Again, the currency in which the series are expressed matters for the interpretation and especially for the evaluation of statistical significance. When expressed in €, we see that GDP growth reacts negatively following an oil supply shock. The effect becomes and remains statistically significant after 8 quarters. Expressed in $ however, we cannot identify any significant effect of an oil supply shock. Regarding the reaction following an oil price increase due to an aggregate demand shock, we see no significant reaction of German GDP growth regardless

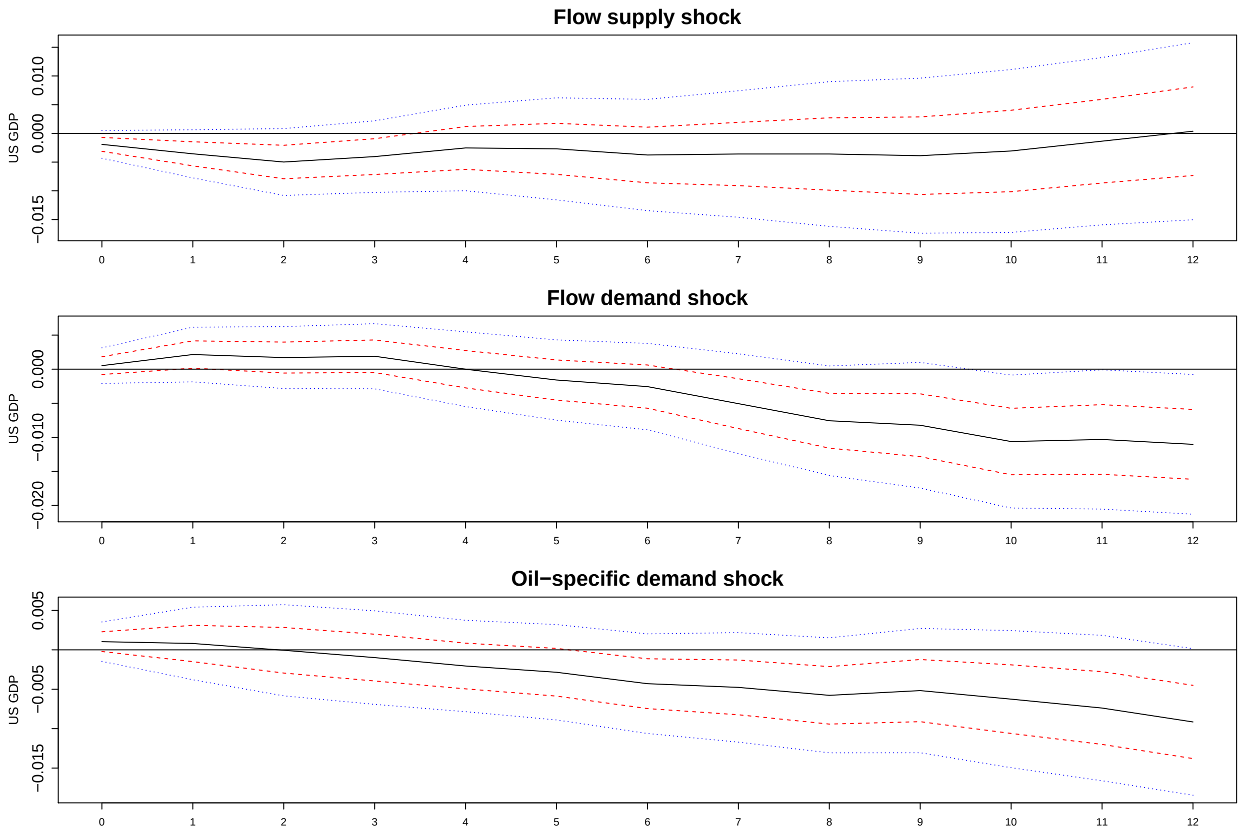

Figure 5.10: Cumulated responses of US GDP growth to each structural shock. (Point estimates with one- and two-standard error bands)

whether the series is expressed in € or \$. As for the influence of an oil-specific demand shock on German GDP growth, we see no significant effect when the series is expressed in €. When expressed in \$ however, we identify a temporary negative and significant effect from quarter 2 until quarter 7 after the shock.

Finally, figure 5.10 shows the cumulated responses of US GDP growth to all three structural shocks. The responses are in line with the original results presented in Kilian (2009, p. 1067) regarding US GDP. Oil supply disruptions have no significant effect on US GDP growth. On the other hand, oil price disruptions caused by a global aggregate demand shock or an oil-specific demand shock have a significant negative growth effect after a delay of around 8 quarters.

In tables 5.2 and 5.3 we show the test results of the Wald χ^2-test regarding the cumulative significance on the monthly and quarterly series respectively. In addition to the p-value, the significance levels are expressed by the usual stars. The result that not all variables are

		$H_0 : \Phi_0 + \ldots + \Phi_k = 0$								
		$k = 12$			$k = 24$			$k = 36$		
		χ^2		p	χ^2		p	χ^2		p
Oil supply shock	Ger Exp (€)	0.001		0.971	0.67		0.413	1.733		0.188
	Ger Imp (€)	0.128		0.721	1.302		0.254	2.907	*	0.088
	Ger Exp ($)	0.408		0.523	1.801		0.18	2.516		0.113
	Ger Imp ($)	0.629		0.428	2.71		0.1	4.189	**	0.041
	Ger IP	1.145		0.285	3.294	*	0.07	4.69	**	0.03
	Ger UN	14.962	***	0	21.051	***	0	34.291	***	0
	Ger CPI	12.856	***	0	13.977	***	0	10.754	**	0.001
	US Exp	1.129		0.288	5.092	**	0.024	5.548	**	0.019
	US Imp	0.556		0.456	0.024		0.877	0.021		0.886
	US IP	0.02		0.887	0.631		0.427	0.583		0.445
	US UN	4.106	*	0.043	6.854	**	0.009	5.008	*	0.025
	US CPI	3.94	**	0.047	2.082		0.149	0.53		0.466
	$/€	0.466		0.495	0.848		0.357	0.861		0.353
Aggregate demand shock	Ger Exp (€)	3.227	**	0.072	0.488		0.485	0.102		0.75
	Ger Imp (€)	3.989	**	0.046	1.432		0.231	0.31		0.577
	Ger Exp ($)	1.098		0.295	0.025		0.875	0.027		0.869
	Ger Imp ($)	1.461		0.227	0.288		0.591	0.001		0.982
	Ger IP	4.996	**	0.025	0.118		0.731	0.01		0.921
	Ger UN	1.945		0.163	0.063		0.801	0.322		0.571
	Ger CPI	8.164	**	0.004	6.977	**	0.008	5.771	**	0.016
	US Exp	10.106	**	0.001	3.009	*	0.083	0.987		0.32
	US Imp	4.09	**	0.043	0.014		0.906	0.265		0.607
	US IP	0.321		0.571	4.596	**	0.032	5.165	**	0.023
	US UN	0.051		0.822	4.871	**	0.027	7.804	**	0.005
	US CPI	8.956	**	0.003	12.961	***	0	9.226	**	0.002
	$/€	0.029		0.866	0.057		0.811	0.15		0.699
Oil-specific demand shock	Ger Exp (€)	2.343		0.126	0.595		0.441	0.005		0.945
	Ger Imp (€)	5.49	**	0.019	0.294		0.588	0.14		0.708
	Ger Exp ($)	0.027		0.87	0.058		0.81	0.333		0.564
	Ger Imp ($)	0.89		0.345	0.118		0.731	0.708		0.4
	Ger IP	2.397		0.122	0.449		0.503	0.004		0.947
	Ger UN	0.658		0.417	0.664		0.415	2.428		0.119
	Ger CPI	6.722	**	0.01	6.517	**	0.011	5.669	**	0.017
	US Exp	0.952		0.329	1.408		0.235	3.921	**	0.048
	US Imp	0.908		0.341	0.845		0.358	0.896		0.344
	US IP	0.019		0.891	1.03		0.31	2.753	*	0.097
	US UN	0.019		0.889	1.33		0.249	1.192		0.275
	US CPI	14.757	***	0	7.977	**	0.005	4.757	**	0.029
	$/€	1.021		0.312	0.613		0.434	0.463		0.496

Signif. codes: $0 <$ *** $< 0.001 <$ ** $< 0.05 <$ * < 0.01

Table 5.2: Test results regarding the monthly cumulative significance of all three structural shocks up to 12, 24 and 36 months.

$H_0 : \Phi_0 + \ldots + \Phi_q = 0$

		$q = 12$			$q = 24$			$q = 36$		
		χ^2		p	χ^2		p	χ^2		p
Oil supply demand shock	Ger (€)	3.737	*	0.053	4.866	**	0.027	5.888	**	0.015
	Ger ($)	0.26		0.61	1.367		0.242	2.76	*	0.097
	US	0.461		0.497	0.324		0.569	0.002		0.961
Aggregate demand shock	Ger (€)	0.023		0.88	1.371		0.242	1.184		0.277
	Ger ($)	0.597		0.44	0.024		0.876	0.061		0.805
	US	0		0.997	3.549	**	0.06	4.632	**	0.031
Oil-specific demand shock	Ger (€)	0.836		0.36	0.984		0.321	2.243		0.134
	Ger ($)	1.69		0.194	1.397		0.237	0.465		0.496
	US	0.5		0.479	2.505		0.114	3.875	***	0.049

Signif. codes: $0 <$ *** $< 0.001 <$ ** $< 0.05 <$ * < 0.01

Table 5.3: Test results regarding the quarterly cumulative significance of all three structural shocks up to 4, 8 and 12 quarters on German and US GDP.

affected by the three structural oil price shocks seems to be confirmed here. Furthermore, as in the case of the IRFs, we observe temporary significant effects over the evaluation horizon. We will now review the main results regarding the German and US reactions to each shock. We point out, that we evaluate the signs based on the results from the IRFs discussions.

With regard to an oil supply shock, the test results confirm the main insights mentioned. German unemployment growth and CPI inflation react significantly over the response horizon and are higher after. The US unemployment growth reacts similarly but weaker from a statistical significance point of view. The quarterly GDP series in € for Germany reacts significantly and negatively after 12 months while no effect is observed for the US GDP series.

With regard to oil price changes that result from a positive aggregate demand shock for industrial commodities, we observe the same positive significant and temporary effect on German exports, imports (expressed in €) and industrial production. German CPI inflation growth is higher following an aggregate demand shock and the effect is significant on the 5% level over the 36 months. No effect on German GDP is observed. In the US the effects are different. We confirm that US economy is negatively affected by an aggregate demand shock in the long run overshadowing the short run positive effects from strong global activity. Indeed after an aggregate demand shock, industrial production and GDP growth are smaller while unemployment growth and CPI inflation are stronger. All effects are statistically significant between the 1% and 5% significance levels.

Finally, the effects of an oil-specific demand shock can be summarized as follows. German imports show a significant temporary increase when expressed in € over the first 12 months. While the effect on German GDP growth becomes weakly significant and negative in the last 12 months of the evaluation horizon. The US economy on the other hand is affected by oil-specific demand shocks in two primary ways. First, we see that CPI growth is higher over the 36 months evaluation horizon. Second, US GDP growth becomes negatively affected after four quarters and the effect remains statistically significant until the end of the evaluation period.

5.3 Summary and discussion

The purpose of the analysis was to evaluate the effect of the three types of structural oil price shocks on different monthly and quarterly macroeconomic series for Germany and the US. We confirm the primary result of Kilian (2009) in that there are differences in the effect of an oil price shock when the underlying causes are differentiated as modeled by the identifying SVAR restrictions. Furthermore we find that the two developed economies of Germany and the United States, namely Germany and the United States, react differently to an oil supply shock, an aggregate demand shock or an oil-specific demand shock.

The German economy is primarily affected negatively by oil price shocks that result from adverse supply disruptions. Indeed, we see that the industrial production and GDP growth rates are significantly lower in the 36 months following a supply shock. Similarly, unemployment growth and CPI inflation are significantly higher after an oil supply shock. Weaker adverse effects on CPI inflation and GDP growth are observed following an oil-specific demand shock. In the case of an aggregate demand shock for global industrial commodities, we even see that there are temporary positive effects on the economy. Export, import and industrial production growth are temporarily higher, while unemployment growth is temporarily significantly lower in the aftermath of an aggregate demand shock. Price pressures remain insignificant over the evaluation period.

The US on the other hand is mostly affected by the adverse effects in the aftermath of oil price increases resulting from aggregate demand and oil-specific demand shocks. Industrial production and GDP growth rates are affected adversely after an initial positive reaction following an aggregate demand shock. The same effect is observed on unemployment. The initial positive effect on the US labor market becomes significantly negative and persistent, resulting in higher unemployment growth rates. CPI inflation is also significantly higher after an aggregate demand shock. Oil-specific demand shocks affect the US economy primarily by lowering GDP growth rates. Surprisingly, the classical notion of an oil price shock resulting from supply disruptions has surprisingly weak if any effects on the US economy. Only unemployment growth seems to be adversely affected.

Recalling from figure 4.9 in section 4.4.2 from the previous chapter, that historically, aggregate demand shocks and oil-specific demand shocks were more frequent and impactful in explaining important oil price movements, it becomes easier to understand why the research has primarily highlighted the adverse effects of oil price changes when using US data. Here we confirm the main results from Kilian (2009): Not supply shocks are the main cause of macroeconomic disruptions for the US when faced with higher oil prices, but more importantly higher prices because of a strong global demand for industrial commodities and to a lesser extend adverse oil-specific demand shocks. For the US economy, the positive effects that are induced by a strong global economy dissipate over time and the negative effects resulting from higher oil prices prevail.

Germany on the other hand, reacts negatively primarily to oil price increases due to physical supply disruptions. Again, given that these were observed less frequently in our sample from January 1974 to December 2017, this might explain the general robustness of the German economy when we only observe rising oil prices without differentiating the underlying causes. It furthermore helps to understand previous surprising empirical results regarding the reaction of the German economy to oil price changes. Indeed, Kilian (2008a, p. 94) is puzzled when estimating a positive effect on GDP growth in the first two quarters after the German economy is faced with a measure of an exogenous oil price shock as constructed in Kilian (2008b). Similarly Blanchard and Gali (2007, p. 19) find that for some countries, including Germany, Italy and Japan, the macroeconomic responses to higher oil prices "fit conventional wisdom less well". Indeed they use nominal oil price changes as a measure of an oil price shock, not disentangling the underlying causes as proposed by the structural VAR framework. As we saw, German macroeconomic aggregates react positively when the cause of an oil price increase in a global aggregate demand shock for industrial commodities or an oil specific demand shock, both shock types were primarily observed in our evaluation sample in explaining important oil price shifts.

A natural follow up question that might be answered by subsequent research regards the explanation of these differences in the macroeconomic effects of the structural oil price shocks.

Including more countries in an analysis and clustering them into similar groups regarding structural characteristics such as for example the degree of dependency on foreign crude oil imports, energy taxes, energy efficiency policies and GDP share of international trade might allow to identify possible explanations. Furthermore the theoretical and empirical literature, that up until now relied on the observed oil price changes without decomposing the underlying contribution of the different structural causes, offers important insights into how channels of transmissions are affected by the different shocks. How the three structural oil price shocks affect consumer expeditures (see Edelstein and Kilian, 2007a, 2009) or investment decisions (Edelstein and Kilian, 2007b; Hamilton, 2016) or prices and monetary policy (see Bernanke et al., 1997; Barsky and Kilian, 2002; Hamilton and Herrera, 2004) are questions that need to be reviewed given these new methodological advances.

In the next chapter we apply the SVAR models presented in chapter 4 to a framework to evaluate the impact of the exogenous sanctions imposed by the US and the EU on the Iranian economy on the global price of crude oil in the years 2011/12 and late 2018.

Chapter 6

Iran sanctions and global oil prices: structural model comparison and forecast evaluation

An important policy implication of the political and economic sanctions imposed by the US and the EU on Iran might be direct changes regarding the global price of crude oil . If, for example, sanctions result in global oil prices increases, it might pose a threat to the global economy, or more specifically, to oil importing countries and consumers. On the other hand, countries that heavily rely on crude oil exports might benefit from sanctions, even if they are not directly involved. The Iranian sanctions, thus offer an interesting framework to evaluate and compare the forecast performance of both presented SVAR models. The first one to evaluate will be the three-variable SVAR model from Kilian (2009) that is recursively identified as described in section 4.2.1. The second model is the four-variable SVAR model from Kilian and Murphy (2014) that is identified by sign restrictions as well as oil market specific conditions, as described in section 4.2.2.

The goal of this chapter is to primarily use the structural framework previously estimated and presented to empirically quantify the impact of the 2011/12 sanctions on the global price of oil, however, sanctions and the economic mechanisms, through which they may be successful, remain hardly understood by economists (see Eaton and Engers, 1992, 1999). Empirically, sanctions have been mostly tested using the gravity framework of trade, but results suggest that they seem to be ineffective and generally do not succeed in achieving their goals (see Hufbauer et al., 2009, for an overview of research). While economic sanctions are indeed harmful to GDP growth of the targeted country (Neuenkirch and Neumeier,

This chapter is based on joint work with Sarah Mohring.

2015), the primary goals that aim at policy or even regime changes mostly fail (Cortright and Lopez, 2000). Additionally, economic sanctions seem to induce adverse effects not at the center of policymakers. Sanctions reduce the targeted government's protection for human right (Peksen, 2009). The institutional level of democracy in a country is also negatively impacted by sanctions (Peksen and Drury, 2010). In the current international debate, the cases of Iran or North Korea are good example of sanctions that failed to reach their political goals.

This chapter is organized as follows: The first section provides a brief overview over the Iranian oil sector as well as the international sanctions. The sanctions imposed in 2011 and lifted in 2014 with the implementation of the Joint Comprehensive Plan of Action shall be described. The idea is to evaluate the forecast accuracy of both SVAR models when introducing exogenous supply shocks to the global crude oil production that correspond to the decrease and later increase in observed Iranian crude oil production. Methodologically, the construction of conditional forecasts that rely on sequences of structural shocks similar to Baumeister and Kilian (2014a) are used. Based on the evaluation results, estimates are presented on how the global oil price would have developed in the absence of the 2011/12 sanctions to calculate the direct effect of the Iran sanctions. These are used on the real price of oil. A counterfactual outcome for each model (Kilian and Lütkepohl, 2017, pp. 128-34) is created by simulating a different path of the VAR variables under the assumption that the structural flow supply shocks differ from their counterparts estimated from the observed data. The impact of the renewed US sanctions since the withdrawal from the Joint Comprehensive Plan of Action (JCPOA) will then be qualitatively evaluated with view to their model characteristics as well as the discussions presented in chapter 4 with respect to the historical influence of the structural shocks. The chapter concludes with a summary of the main methodological findings as well as on the effects of sanctions on the price of oil.

6.1 The Iranian oil sector and the international sanctions

Historically Iran has played an important role in the development of the global oil industry. With the help of British companies, the Iranian oil industry grew in the 1930s converting the country into one of the major oil producers worldwide Yergin (1991, p. 148-149). As can be seen in figure 6.1, before the Iranian revolution in 1979, the country's crude oil output amounted at around 6 million barrels per day (mbd). The chaos that followed the revolution as well as the start of the Iran-Iraq war in September 1980 resulted in important losses in production capacity that stabilized around 1984 at a daily production of 2 mbd. Only after

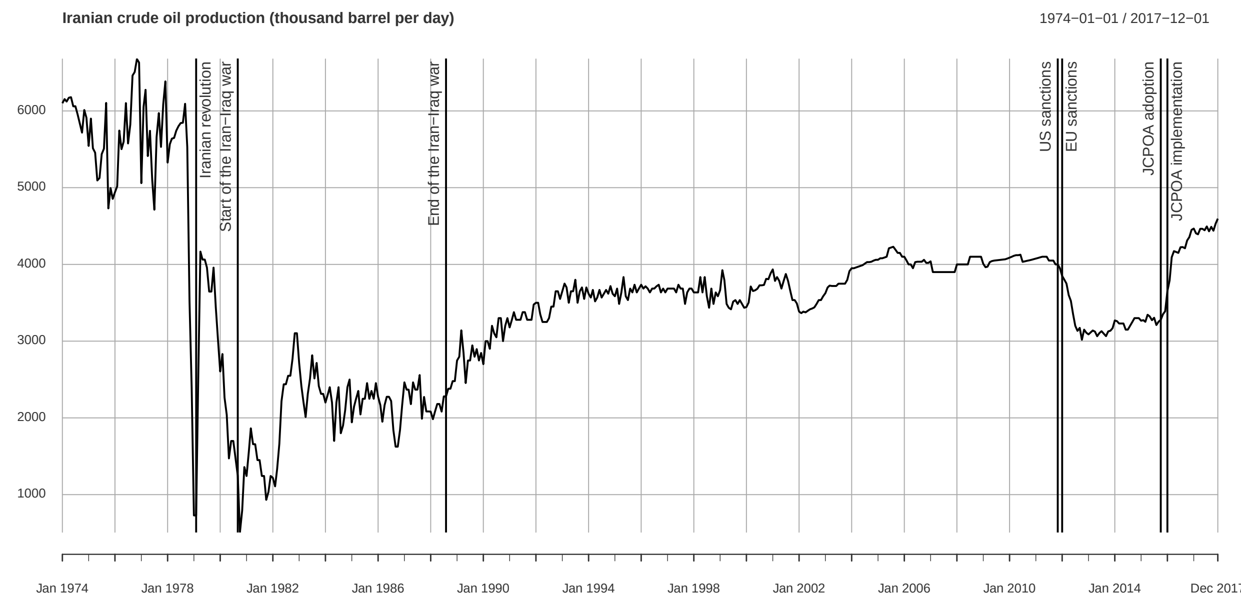

Figure 6.1: Evolution of the Iranian crude oil production between January 1974 and December 2017 (Source: EIA).

the first Gulf war, Iranian oil production started to grow, stabilizing at around 4 mbd in 2011. According to the EIA, 2019a Iran holds the fourth largest crude oil reserves as well as the second largest natural gas reserves in the world. The main reason why the Iranian oil industry never recovered to pre-revolution levels are the sanctions, imposed by the international community and especially, the US. Reference is given to Katzman (2019) for an up-to-date, detailed analysis of all sanctions, imposed by the US and the United Nations from 1979 onwards, when President Jimmy Carter used them as instruments on the Country. We now briefly discuss the developments since 2010 specifically targeting the oil and energy sector that are relevant for the assumptions in this analysis.

Suffering from political and economic sanctions since 1979, with an initial focus on technological and military sectors, in 2006 and 2010 additional sanctions also focused on imports of petroleum products. As Van de Graaf (2013) notes, the discovery of the Iranian nuclear program led to an increase in sanctions introduced by the international community through the United Nations in the late 2000's. In 2010, the European Union imposed for the first time its own autonomous sanctions that specifically targeted investments in Iranian oil and gas sectors to further hurt the Iranian economy and to force the country to abandon its nuclear program (Patterson, 2013). The most severe sanctions, however, were imposed in late 2011 and early 2012, when the US and the EU targeted transactions with the Iranian central bank as well as banning European companies from providing insurance and transport services with regard to Iranian oil exports. European imports of Iranian crude oil were also banned, which in 2011 still accounted for around one quarter of Iran crude oil exports (Patterson, 2013, p. 153). The decline in crude oil production resulting from these specific oil-related sanctions can be clearly seen in figure 6.1.

In 2015, representatives from France, China, Russia, the United Kingdom, the United States, Germany, the EU and Iran signed the JCPOA agreement in Vienna. The deal aimed at constraining the Iranian nuclear program to be strictly monitored by international inspectors, but also promising Iran lifting previously imposed nuclear-related sanctions (European Council, 2019). For Iran it provided the opportunity to attract foreign investment into its old and inefficient oil sector. However, in May 2018 the United States under President Trump unilaterally withdrew from the JCPOA and formally imposed new sanctions in November 2018 (Katzman, 2019). These international developments can be detected from declining oil production and export figures for Iran, which originally started to improve after signing the JCPOA (EIA, 2019a).

6.2 Structural model comparison and forecast evaluation

In this section, the in-sample forecast accuracy of the presented structural models of chapter 4 will be evaluated, including exogenous flow supply shock series that result from the sanctions on Iran. It is useful to look at figure 6.2 that shows the evolution of the Iranian crude oil production between January 2010 and December 2017. It seems evident, that the US and EU sanction rounds in late 2011 and early 2012 had an immediate negative impact on the Iranian crude oil output. Under their effect, daily crude oil production fell from around 4 to 3.1 million barrels. This production level started to increase as soon as the JCPOA was announced in October 2015 and also nuclear-related sanctions on the crude oil sector were soon to be lifted. Two different phases can be identified, which characterize the expansion of the Iranian crude oil production after the announcement and implementation of the JCPOA. A rapid increase to pre-2011/12 sanction levels can be observed, followed by a slow and steady expansion of Iranian crude oil production, as the JCPOA also guaranteed the removal of previous sanctions.

Under the assumption of knowing how the Iranian crude oil production would react to the implementation and removal of sanctions, we can construct a series of exogenous crude oil

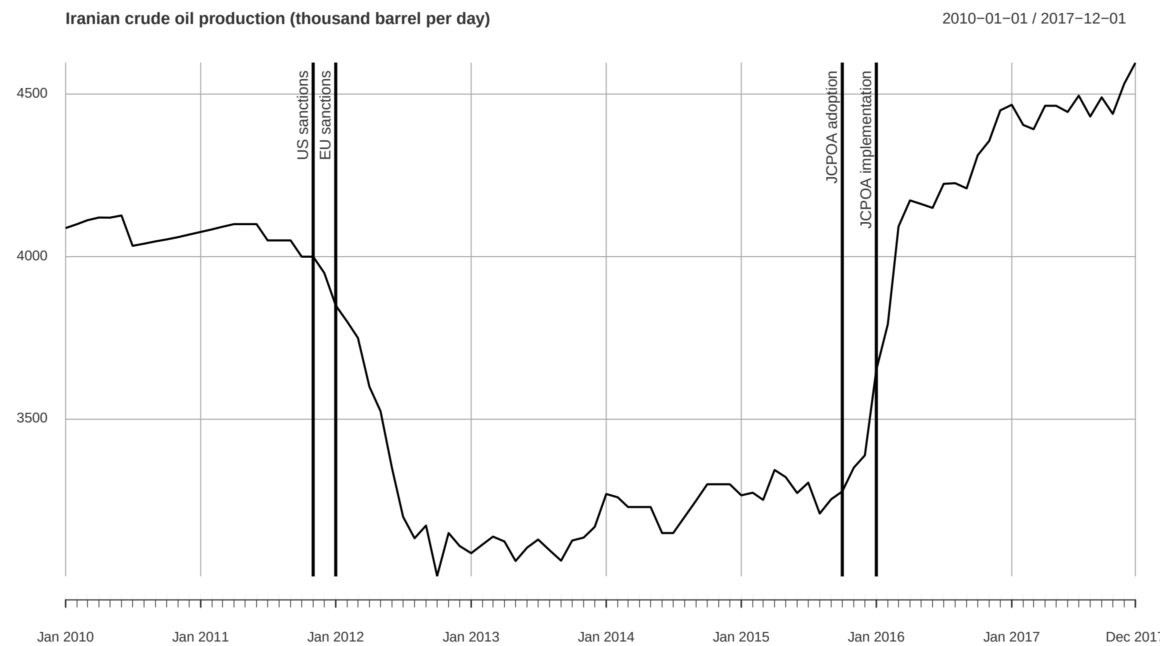

Figure 6.2: Evolution of the Iranian crude oil production between January 2010 and December 2017 (Source: EIA).

supply shocks and introduce them into both global models of oil for evaluation purposes. Here we content ourselves with the simple assumption, that the cumulative production changes due to the sanctions are equal to the difference between the observed pre-sanction and the observed post-sanction levels. In other words, we assume that the reduction in Iranian oil production that after November 2011 is solely the consequence of the imposed sanctions. The same inverse assumption is held true for the removal of sanction in October 2015. Consequently, the first cumulative production shock series is negative and the second cumulative shock series is positive.

The month-to-month changes in this difference in levels is called $\Delta iran_t$ and can be seen in figure 6.3. The left panel indicates, that the observed effects of the sanctions imposed in 2011/12 occur primarily during the first eight months. Afterwards, positive and negative changes mostly compensate each other on a month-to-month basis. The right panel indicates similar, but faster effects implied by the removal of sanctions as agreed by the JCPOA. Within the first five months, most previously lost crude oil production is recovered and a weak positive trend can be identified.

For both, the imposition of sanctions in 2011/12 and their removal in late 2015, the per-

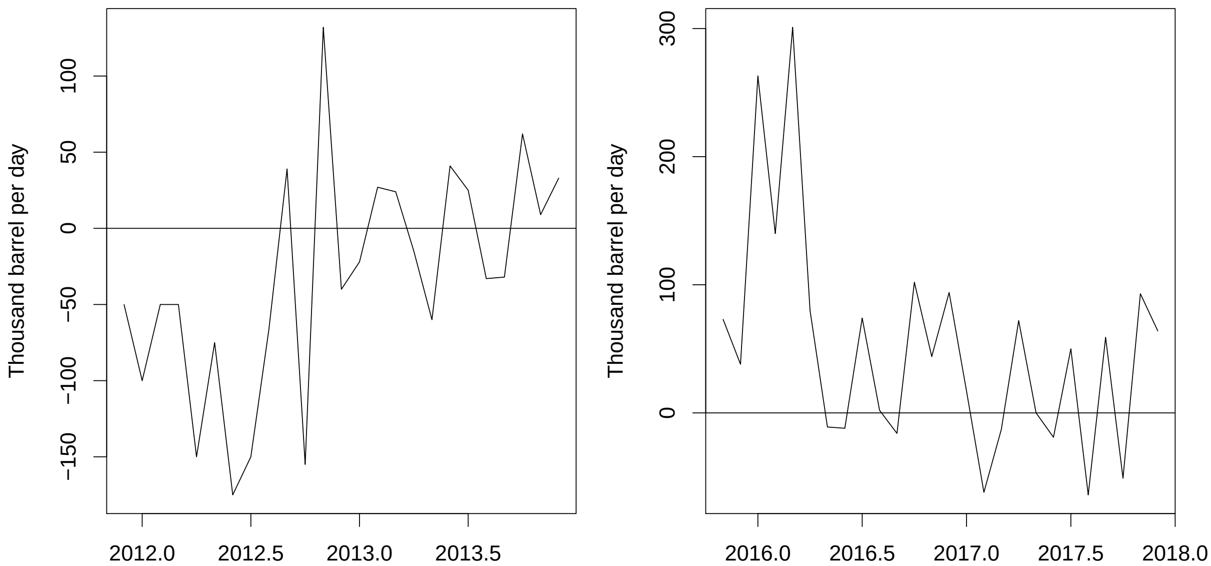

Figure 6.3: Exogenous shock series $\Delta iran_t$ for the 2011/12 (left) and 2015/16 (right) implementation and removal of sanctions (Source: Own calculations based on EIA).

spective of a forecaster is adopted who observes the VAR data up until the point in time, in which sanctions were imposed (removed) in November 2011 (October 2015). The forecaster then proceeds to recursively forecast out-of-sample up until December 2013, December 2017 respectively. Goal is to create a "what, if" analysis given that the information contained in the Iranian supply shock series. In comparison to simple out-of-sample VAR forecasts, that as Kilian and Lütkepohl (2017, p. 120-21) point out, remains the best out-of-sample forecast based on past information, we thus include one structural flow supply shock and, by transformation, all reduced form shocks are different from zero. This allows to evaluate the performance of both structural models by comparing their "what if" forecasts with the observed data that also includes the reduced Iranian oil output. In reality, such an out-of-sample estimate would first require a forecaster to assess the impact of the sanctions on Iranian crude oil output to then include this information in his models.

As a reminder from section 4.3, the global crude oil production variable is expressed in the differences of the logarithms for both structural models. A transformation of monthly exogenous shocks into the same format is required before being able to properly assess the forecast characteristics of both models. This requires to iteratively forecast all endogenous variables starting from period $t_0 + 1$ that follows period t_0 in which the sanctions are implemented or removed. Thus, for the first evaluation t_0 is set to November 2011 and the first forecast is done for December 2011. For the second evaluation t_0 is set to October 2015 and the first forecast is made for November 2015. We start by forecasting all VAR variables $\widehat{\boldsymbol{y}}_{t_0+1}$ given that no structural shocks occur by using the following reduced form VAR equation[1]:

$$\widehat{\boldsymbol{y}}_{t_0+1} = \Sigma_{i=1}^{24} \widehat{\mathbf{A}}_i \boldsymbol{y}_{t_0+1-i}. \tag{6.1}$$

As crude oil production for t_0 is known, one can calculate the crude oil production (expressed in thousand barrels per day) for period $t_0 + 1$ called $\widehat{prod}_{t_0+1}$ given the estimate of the first VAR variable $\widehat{y}_{1(t_0+1)}$:

$$\widehat{y}_{1(t_0+1)} = \ln\left(\frac{\widehat{prod}_{t_0+1}}{prod_{t_0}}\right) \qquad \Rightarrow \qquad \widehat{prod}_{t_0+1} = prod_{t_0} \exp\left(\widehat{y}_{1(t_0+1)}\right). \tag{6.2}$$

Knowing the level of the implied exogenous shock $\Delta iran_{t_0+1}$ it is now possible to define an alternative level of global production that includes the Iran sanctions:

$$\widehat{prod}_{t_0+1}^{with\,Iran} = \widehat{prod}_{t_0+1} + \Delta iran_{t_0+1}$$

It is now possible to correctly construct the first element in the reduced form residual vector $\widehat{u}_{t_0+1}$ that we defined as $\widehat{u}_{t_0+1}^{\Delta prod}$ by using the following transformation:

$$\begin{aligned}
& \ln\left(\frac{\widehat{prod}_{t_0+1}}{prod_{t_0}}\right) + \hat{u}_{t_0+1}^{\Delta prod} && = \ln\left(\frac{\widehat{prod}_{t_0+1}^{with\,Iran}}{prod_{t_0}}\right) \\
\Longleftrightarrow \quad & \hat{u}_{t_0+1}^{\Delta prod} && = \ln\left(\frac{\widehat{prod}_{t_0+1}^{with\,Iran}}{prod_{t_0}} \frac{prod_{t_0}}{\widehat{prod}_{t_0+1}}\right) \\
\Longleftrightarrow \quad & \hat{u}_{t_0+1}^{\Delta prod} && = \ln\left(\frac{\widehat{prod}_{t_0+1}^{with\,Iran}}{\widehat{prod}_{t_0+1}}\right) \\
\Longleftrightarrow \quad & \hat{u}_{t_0+1}^{\Delta prod} && = \ln\left(1 + \frac{\Delta iran_{t+1}}{\widehat{prod}_{t_0+1}}\right)
\end{aligned}$$

Next the whole vector of reduced form errors $\widehat{\boldsymbol{u}}_{t_0+1}$ is constructed, using its relationship with the structural errors $\widehat{\boldsymbol{\varepsilon}}_{t_0+1}$ with the help of the estimated impact multiplier matrix $\widehat{\boldsymbol{B}}$. In the following the construction of the four-variable SVAR model is shown. The construction for the 3-variable SVAR model is equivalent with omitting the fourth variable regarding above ground crude oil inventories. Since the change in Iranian crude oil production is the sole

[1] Recall that for notational purposes we suppressed the seasonal dummies that are included in the model.

exogenous structural shock only $\widehat{\varepsilon}_{t_0+1}^{flow\,supply\,shock}$ is assumed to be different from zero.

$$\hat{\boldsymbol{u}}_{t_0+1} \equiv \begin{pmatrix} \hat{u}_{t_0+1}^{\Delta prod} \\ \hat{u}_{t_0+1}^{rea} \\ \hat{u}_{t_0+1}^{rpo} \\ \hat{u}_{t_0+1}^{\Delta inv} \end{pmatrix} = \begin{bmatrix} \hat{b}_{11} & \hat{b}_{12} & \hat{b}_{13} & \hat{b}_{14} \\ \hat{b}_{21} & \hat{b}_{22} & \hat{b}_{23} & \hat{b}_{24} \\ \hat{b}_{31} & \hat{b}_{32} & \hat{b}_{33} & \hat{b}_{34} \\ \hat{b}_{41} & \hat{b}_{42} & \hat{b}_{43} & \hat{b}_{44} \end{bmatrix} \begin{pmatrix} \widehat{\varepsilon}_{t_0+1}^{flow\,supply\,shock} \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{6.3}$$

Taking $\hat{u}_{t_0+1}^{\Delta prod}$ and $\hat{\boldsymbol{B}}$ from the SVAR estimate, we are able to compute $\widehat{\varepsilon}_{t_0+1}^{flow\,supply\,shock}$ and thus the other components of $\hat{u}_{t_0+1}$.

$$\begin{aligned} \hat{u}_{t_0+1}^{\Delta prod} &= \hat{b}_{11}\ \widehat{\varepsilon}_{t_0+1}^{flow\,supply\,shock} \qquad \Leftrightarrow \qquad \widehat{\varepsilon}_{t_0+1}^{flow\,supply\,shock} = \hat{u}_{t_0+1}^{\Delta prod}/\hat{b}_{11} \\ \hat{u}_{t_0+1}^{rea} &= \hat{b}_{21}\ \widehat{\varepsilon}_{t_0+1}^{flow\,supply\,shock} \\ \hat{u}_{t_0+1}^{rpo} &= \hat{b}_{31}\ \widehat{\varepsilon}_{t_0+1}^{flow\,supply\,shock} \\ \hat{u}_{t+1}^{\Delta inv} &= \hat{b}_{41}\ \widehat{\varepsilon}_{t_0+1}^{flow\,supply\,shock} \end{aligned}$$

It is now possible to construct the forecasts of all VAR variables for $t_0 + 1$ controlling for the effects of the exogenous oil supply shock, as implied by the sanctions imposed or removed on Iran:

$$\widehat{\boldsymbol{y}}_{t_0+1}^{with\,Iran} = \Sigma_{i=1}^{24} \widehat{\mathbf{A}}_i \boldsymbol{y}_{t_0+1-i} + \widehat{\boldsymbol{u}}_{t_0+1} \tag{6.4}$$

The next step derives itself by inserting the newly constructed value of the endogenous variable vector $\widehat{\boldsymbol{y}}_{t_0+1}^{with\,Iran}$ into equation (6.1). This is again first required in order to estimate the reduced form forecast $\widehat{\boldsymbol{y}}_{t_0+2}$ for period $t_0 + 2$:

$$\widehat{\boldsymbol{y}}_{t_0+2} = \widehat{\mathbf{A}}_1 \widehat{\boldsymbol{y}}_{t_0+1}^{with\,Iran} + \Sigma_{i=2}^{24} \widehat{\mathbf{A}}_i \boldsymbol{y}_{t_0+2-i} \tag{6.5}$$

It can be noted that in comparison with equation (6.1) the last required vector, i.e. $\widehat{\boldsymbol{y}}_{t_0+1}^{with\,Iran}$ is not a sample observation, but the first estimate, assuming that the Iranian supply shock is included. As further steps ahead are estimated, the required number of sample observations is reduced, whereas the needed count of estimates increases. After estimating the next reduced form forecast, it is again possible to reconstruct the forecasted level of global crude oil production (expressed in thousand barrels per day) in period $t_0 + 2$ assuming that no shock occurs $\widehat{prod}_{t_0+2} = \widehat{prod}_{t_0+1} \exp\left(\widehat{y}_{1(t_0+2)}\right)$. This value is then increased by the assumed Iranian production shock $\Delta iran_{t_0+2}$ (in thousand barrels per day) in order to calculate the reduced form error $\hat{u}_{t_0+2}^{\Delta prod}$ that is the first element of the reduced form error vector $\hat{\boldsymbol{u}}_{t_0+2}$ to be constructed, following $\hat{u}_{t_0+2}^{\Delta prod} = \ln\left(1 + \frac{\Delta iran_{t+2}}{\widehat{prod}_{t_0+2}}\right)$. Knowing $\hat{u}_{t_0+2}^{\Delta prod}$ and using the coefficients of the estimated impact multiplier matrix $\widehat{\boldsymbol{B}}$, we first construct the exogenous flow supply shock implied by the fall in Iranian production $\widehat{\varepsilon}^{flow\,supply\,shock}_{t_0+2}$ to be able to construct the remaining reduced form shocks that together compose the reduced form vector $\widehat{\boldsymbol{u}}_{t_0+2}$:

$$
\begin{aligned}
\hat{u}^{\Delta prod}_{t_0+2} &= \hat{b}_{11}\ \widehat{\varepsilon}^{flow\,supply\,shock}_{t_0+2} \qquad \Leftrightarrow \qquad \widehat{\varepsilon}^{flow\,supply\,shock}_{t_0+2} = \hat{u}^{\Delta prod}_{t_0+2}/\hat{b}_{11}\\
\hat{u}^{rea}_{t_0+2} &= \hat{b}_{21}\ \widehat{\varepsilon}^{flow\,supply\,shock}_{t_0+2}\\
\hat{u}^{rpo}_{t_0+2} &= \hat{b}_{31}\ \widehat{\varepsilon}^{flow\,supply\,shock}_{t_0+2}\\
\hat{u}^{\Delta inv}_{t+2} &= \hat{b}_{41}\ \widehat{\varepsilon}^{flow\,supply\,shock}_{t_0+2}
\end{aligned}
$$

Finally we construct the forecast of all endogenous variables for period $t = t_0 + 2$:

$$\widehat{\boldsymbol{y}}^{with\,Iran}_{t_0+2} = \widehat{\mathbf{A}}_1 \widehat{\boldsymbol{y}}^{with\,Iran}_{t_0+1} + \Sigma^{24}_{i=2} \widehat{\mathbf{A}}_i \boldsymbol{y}_{t_0+2-i} + \widehat{\boldsymbol{u}}_{t_0+2} \tag{6.6}$$

The same steps are then repeated for all desired forecast periods $t_0 + i$ for $i \geq 2$. As known from section 4.3, the estimated model variables finally require re-transformation in order to be properly illustrated in levels. In the following, the forecast of the nominal price of oil is of primary interest assuming that sanctions are imposed and removed. The characteristics of correctly shocked three- and four- variable SVAR models are evaluated and then the forecasts as implied by the mean forecast combination of both models. The mean forecasts of all estimated four-variable SVAR models are also constructed. In order to show how the structural models perform in comparison to the reduced form models, both the three- and four-variable reduced form VAR models shall be illustrated, while falsely shocking the global crude oil production variable and thus eliminating the instantaneous impact of the remaining endogenous variables to an oil supply shock.

6.2.1 Forecast evaluation results for the 2011/12 sanctions

Figure 6.4 shows the observed nominal price of a barrel of crude oil expressed in US$ (TRUE), the forecast made by falsely shocking the first variable in the three variable VAR model (3_VAR_F), the forecast made by the three variable recursively identified SVAR model that includes the Iranian flow supply shock (3_SVAR), the forecast made by falsely shocking the first variable in the four variable VAR model (4_VAR_F), the forecast made by the four variable sign restricted SVAR model that includes the Iranian flow supply shock (4_SVAR), the mean forecast of both SVAR models (MEAN_BOTH) and the mean forecast of all estimated sign restricted SVAR models that include the Iranian flow supply shock (MEAN_SIGN). Further, for all forecasts the mean squared error (MSE) is calculated in order to complement and check the accuracy evaluation. The results over the forecast horizon

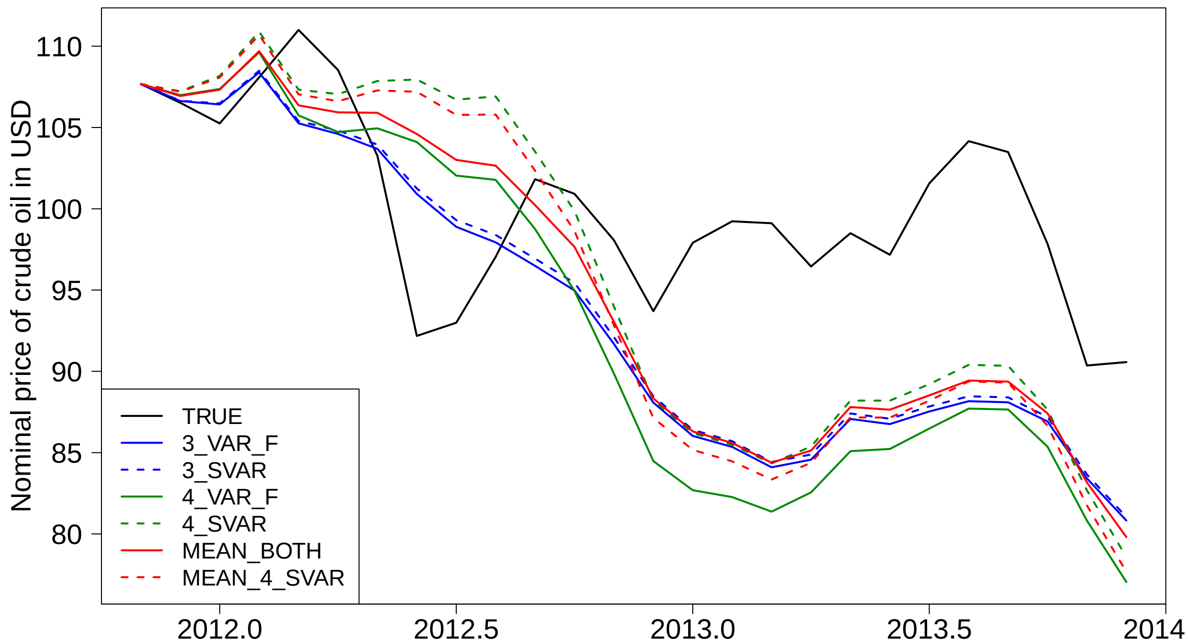

Figure 6.4: Forecasts made by different models including the exogenous oil supply shocks as implied by the 2011/12 sanctions over the forecast horizon December 2011 to December 2013.

are shown in figure 6.5 for all previously discussed models.

Some important insights can be found: First and foremost, the forecasts made with the help of the three-variable models (in blue) beat the forecasts made by the four-variable models (in green). The accuracy divergence between both increases over the first half of the evaluation period. Afterwards, one detects some indications of convergence when looking at the four-variable SVAR model. The three-variable SVAR model remains, however, the most accurate. Furthermore when looking specifically at both forecasts made by the three-variable models (3_VAR_F & 3_SVAR), a difference in accuracy starts to be increasingly pronounced in the second half of the forecast horizon. It primarily indicates, that the information contained in the constructed exogenous shock series is better incorporated by the SVAR model in comparison with the simple reduced form three-variable VAR model. Similarly, when evaluating both four variable models (in green) over the whole forecast horizon, the forecasts made with the SVAR model in comparison with the reduced form VAR model show a smaller MSE. This is, however, not true in the first half of the evaluation sample. Finally, one notes that the forecasts resulting from combinations also beat the four-variable model forecasts.

Concerning the results visualized in figures 6.6 and 6.7 which cover the second evaluation

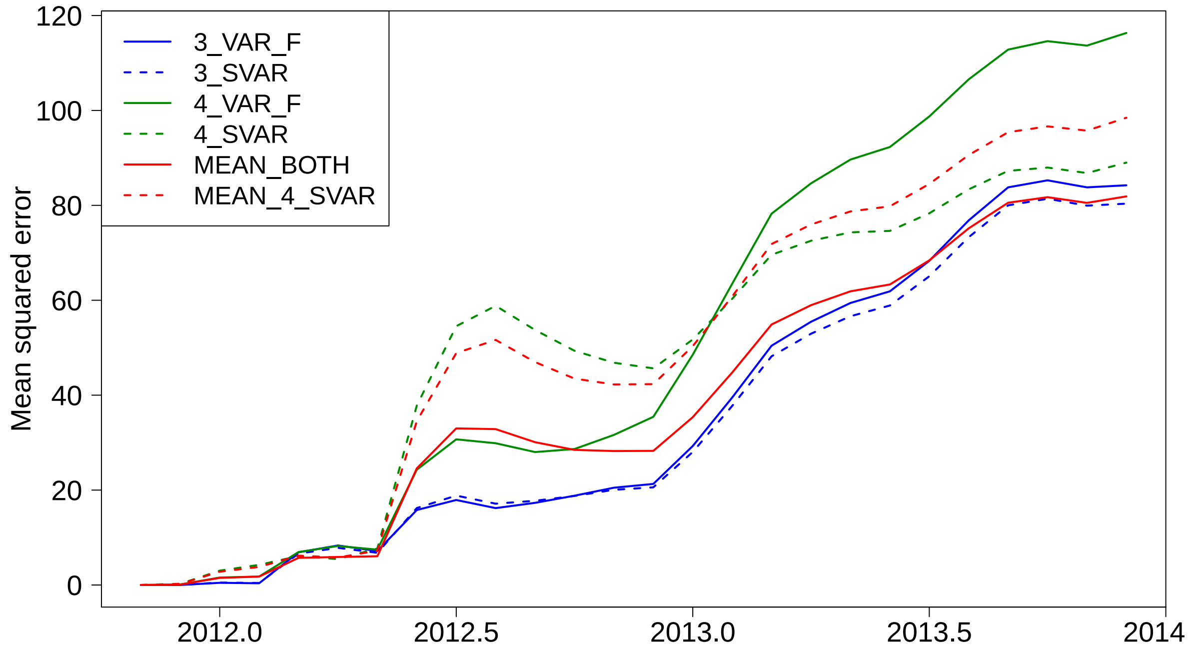

Figure 6.5: Mean squared error over the forecast horizon December 2011 to December 2013.

period after the implementation of the JCPOA and the successive removal of sanctions, the results with respect to accuracy are reversed. The forecasts made by the benchmark four-variable SVAR model as well as forecasts made by combining the results of all estimated four-variable SVAR models are more accurate in comparison with the forecasts made with the help of the three-variable models. In turn, validating the structural assumptions for both the three- and four-variable models, one observes a consistent accuracy increase in comparing the SVAR forecasts with the reduced form VAR forecasts. The accuracy increase is, however, much stronger in the case of the four-variable models.

The conclusion emerges that when analyzing negative exogenous supply shocks the forecasts made using the three-variable SVAR model are the most accurate. On the other hand, when analyzing the impact of positive exogenous supply shocks, the forecasts made with the help of the four-variable SVAR model are more advantageous. Second, the structural assumptions made for both models are validated by the results as the SVAR models are more accurate than their falsely shocked reduced form VAR counterparts. Finally, with respect to forecast combinations concerning all four-variable SVAR forecasts, the results are inconclusive. In the first evaluation sample one observes an increase in accuracy, whereas no such increase could be detected in the second sample where the benchmark four-variable model performed best.

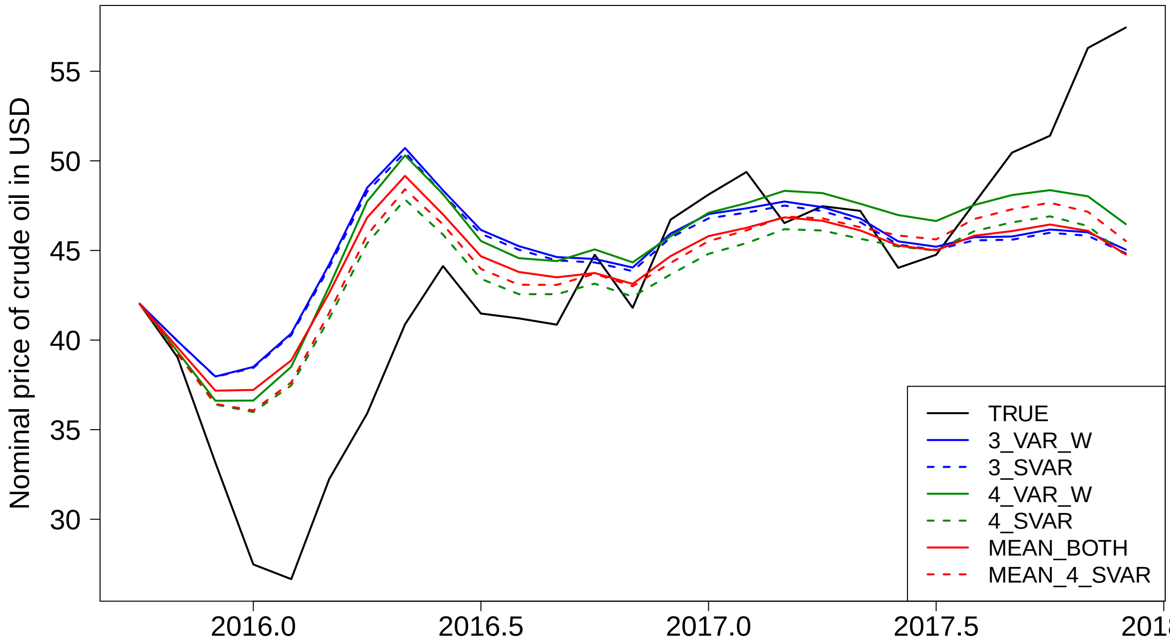

Figure 6.6: Forecasts made by different models including the exogenous oil supply shocks as implied by the JCPOA over the forecast horizon November 2015 to December 2017.

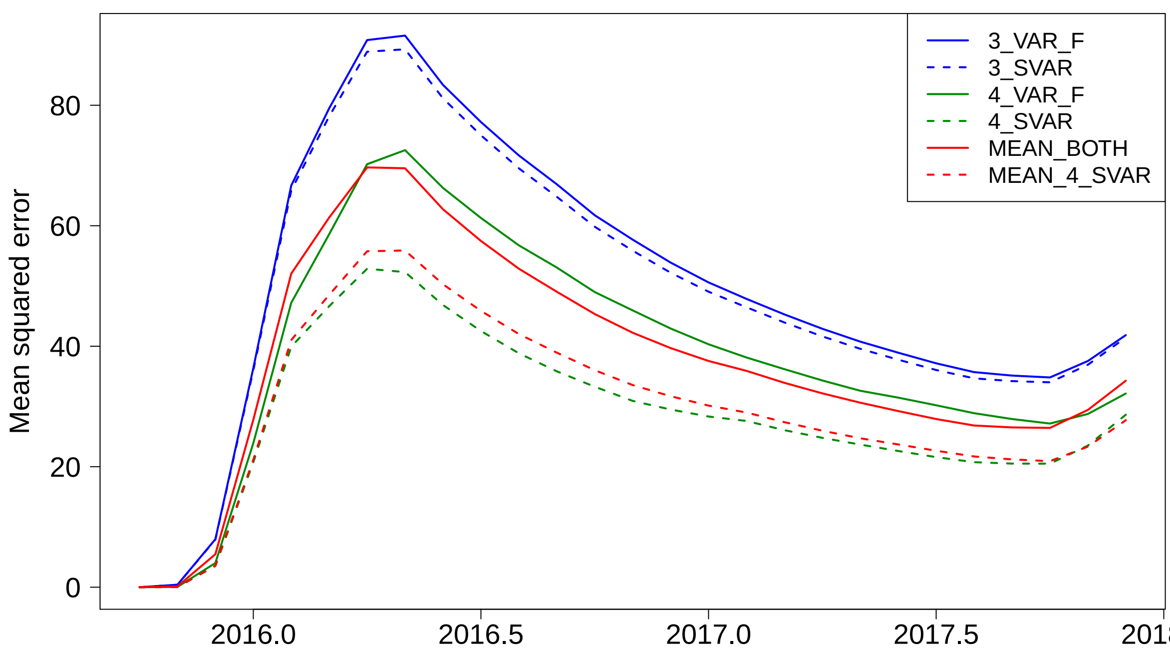

Figure 6.7: Mean squared error over the forecast horizon November 2015 to December 2017.

6.3 How did the sanctions affect the global oil price?

In this section, the effects of sanctions during 2011/2012 on the global price of oil are quantified, using the structural models explained in the previous sections. Furthermore, an attempt is made to qualitatively assess the impact of the US-withdrawal from the JCPOA on the future development of the global market for crude oil and how it might already have affected global oil prices.

6.3.1 Sanctions of 2011/12

The evaluation of the impact of the 2011/12 sanctions on the price of crude oil relies on the construction of counterfactual series for all variables included in the VAR estimation. The analysis is similar to Kilian (2017), who evaluated the impact of the US fracking boom on Arab oil producers by constructing a counterfactual global oil production series, had the US not experienced the fracking boom. He first constructed a counterfactual global crude oil production series, based on assumptions regarding the developments in the United States and their implications on crude oil imports Kilian (2017, p. 144) to then estimate the structural shocks, resulting from the counterfactual production series. Under the assumption that all other structural shocks remain as observed, he was able to determine the *flow supply shocks* required to obtain counterfactual production series with the help of the structural model. he then compared the resulting counterfactual price series with the observed series and attributed the difference to the fracking boom.

Although very similar, the approach proposed for this study requires some modifications, explained in detail in the following: First, the main idea here is to decompose the observed reduced form residuals, primarily the reduced form residual corresponding to the VAR equation concerning the price of crude oil. As it is common knowledge, in reality the Iranian sanctions of 2011/12 were applied and implemented in 2011/12. Because of the fact that they are not modeled by the endogenous variables, the sanction effects are consequently included in the observed reduced form and by consequence structural residuals as implied by both models. Goal in this section is the use of both SVAR models to quantify and remove the effects of negative Iranian supply shocks, which most likely resulted from the imposition of sanctions. The methodology was modified by constructing the counterfactual global production and consequently counterfactual series for all the remaining variables in response to the above mentioned shock series.

The methodology of how to construct counterfactual series is explained for the four-variable model only, the procedure for the three-variable model being identical. The relationship between observations in period t $\boldsymbol{y}_t$, fitted values resulting from the VAR estimate $\widehat{\boldsymbol{y}}_t$ and reduced form residuals $\widehat{\boldsymbol{u}}_t$ is expressed as follows:

$$\boldsymbol{y}_t = \widehat{\boldsymbol{y}}_t + \widehat{\boldsymbol{u}}_t \tag{6.7}$$

The reliance upon a sample period January 1974 to December 2017 to estimate the reduced form VAR model and, consequently, both structural models, determines the values and estimates in equation (6.7) for the sample February 1976 to December 2017. More importantly, one can observe the reduced form residuals $\boldsymbol{u}_t$ and also the structural residuals between November 2011 and December 2013. Within this time frame the counterfactual series for the oil price will be created, assuming that no sanctions were implemented by introducing positive flow supply shocks into the structural models. They correspond with the inverse values that were observed in reality, following the imposition sanctions, shown in figure 6.3. Referring again to equation (6.7), our aim is to modify $\widehat{\boldsymbol{u}}_t$ by removing the effect induced by the Iranian flow supply shock. Similar to the previous section, the first equation regarding the global crude oil production variable, expressed in differences of logarithms of global crude oil production, is taken into consideration as follows:

$$\begin{aligned}
& y_{1t} && = \widehat{y}_{1t} + \widehat{u}_{1t} \\
\iff \quad & \ln\left(\frac{prod_t}{prod_{t-1}}\right) && = \ln\left(\frac{\widehat{prod}_t}{prod_{t-1}}\right) + \hat{u}_t^{\Delta prod} \\
\iff \quad & \frac{prod_t}{\widehat{prod}_t} && = \exp(\hat{u}_t^{\Delta prod})
\end{aligned}$$

$$\iff \qquad prod_t = \widehat{prod}_t \exp(\hat{u}_t^{\Delta prod}) \tag{6.8}$$

$\widehat{prod}_t$, expressed in thousand barrels per day, is easily calculated for every t, satisfying the equation $\widehat{prod}_t = prod_{t-1} \exp(\widehat{y}_{1t})$. Thus we can define the unexpected change in global oil production that is not explained by the endogenous variables as $e_t^{prod} = prod_t - \widehat{prod}_t$. By definition, the unexpected change in global oil production takes into account changes in Iranian oil production that would not occur, if sanctions were not imposed:

$$e_t^{*prod} = e_t^{prod} + \Delta iran_t.$$

In view of the fact, that the fitted value $\widehat{prod}_t$ remains constant, an alternative production series is created, assuming the absence of sanctions $prod_t^* = e_t^{*prod} + \widehat{prod}_t$. Introducing this value for all periods t in equation (6.8) one can calculate the corrected reduced form shock $\hat{u}_t^{*\Delta prod}$, that no longer incorporates the decrease in Iranian crude oil production:

$$\hat{u}_t^{*\Delta prod} = \ln\left(\frac{prod_t^*}{\widehat{prod}_t}\right)$$

Using the coefficients in the estimated impact multiplier matrix $\hat{\boldsymbol{B}}$ this change is attributed to a change in flow supply shocks $\widehat{\varepsilon}_t^{*flow\,supply\,shock}$, whereas all the other three structural shocks $\widehat{\varepsilon}_t^{flow\,demand\,shock}$, $\widehat{\varepsilon}_t^{speculative\,demand\,shock}$, and $\widehat{\varepsilon}_t^{oil\,specific\,demand\,shock}$ remain constant in order to construct the remaining reduced form errors, in the absence of Iranian flow supply shocks:

$$\begin{aligned}
\widehat{\varepsilon}_t^{*flow\,supply\,shock} = \frac{1}{\hat{b}_{11}}(\hat{u}_t^{*\Delta prod} &- \hat{b}_{12}\,\widehat{\varepsilon}_t^{flow\,demand\,shock} \\
&- \hat{b}_{13}\,\widehat{\varepsilon}_t^{speculative\,demand\,shock} - \hat{b}_{14}\,\widehat{\varepsilon}_t^{oil\,specific\,demand\,shock})
\end{aligned}$$

$$\begin{aligned}
\hat{u}_t^{*rea} &= \hat{b}_{21}\,\widehat{\varepsilon}_t^{*flow\,supply\,shock} + \hat{b}_{22}\,\widehat{\varepsilon}_t^{flow\,demand\,shock} \\
&+\ \hat{b}_{23}\,\widehat{\varepsilon}_t^{flow\,supply\,shock} + \hat{b}_{24}\,\widehat{\varepsilon}_t^{oil\,specific\,demand\,shock} \\
\hat{u}_t^{*rpo} &= \hat{b}_{31}\,\widehat{\varepsilon}_t^{*flow\,supply\,shock} + \hat{b}_{32}\,\widehat{\varepsilon}_t^{flow\,demand\,shock} \\
&+\ \hat{b}_{33}\,\widehat{\varepsilon}_t^{flow\,supply\,shock} + \hat{b}_{34}\,\widehat{\varepsilon}_t^{oil\,specific\,demand\,shock} \\
\hat{u}_t^{*inv} &= \hat{b}_{41}\,\widehat{\varepsilon}_t^{*flow\,supply\,shock} + \hat{b}_{42}\,\widehat{\varepsilon}_t^{flow\,demand\,shock} \\
&+\ \hat{b}_{43}\,\widehat{\varepsilon}_t^{flow\,supply\,shock} + \hat{b}_{44}\,\widehat{\varepsilon}_t^{oil\,specific\,demand\,shock}
\end{aligned}$$

In a last step, alternative paths of evolution of endogenous variables $\boldsymbol{y}_t^*$ are constructed, analogous to equation (6.7):

$$\boldsymbol{y}_t^* = \widehat{\boldsymbol{y}}_t + \hat{\boldsymbol{u}}_t^* = \widehat{\boldsymbol{y}}_t + \begin{pmatrix} \hat{u}^{*\Delta prod} \\ \hat{u}_t^{*rea} \\ \hat{u}_t^{*rpo} \\ \hat{u}_t^{*inv} \end{pmatrix}$$

In order to comply with the priority interest in the impact of sanctions on global oil prices, the results of both counterfactual oil price series are explained, assuming that no sanctions were imposed in late 2011 and in early 2012 by the US and the EU. As can be seen in figure 6.8, also incorporating other structural shock estimates into the construction of alternative production series, facilitates smoother predictions, that are much closer to the observed nominal price of oil. This indicates in particular the impact and importance of unexpected shocks

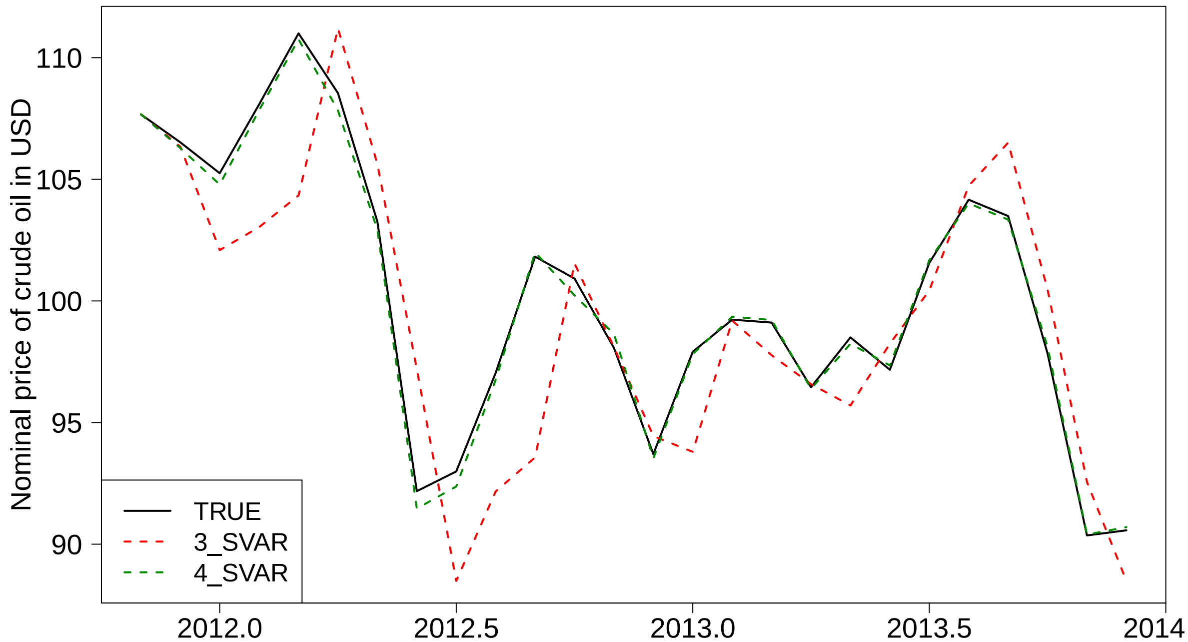

Figure 6.8: Estimated price evolution in alternate scenarios without the Iran sanctions of 2011/12.

on all variables in both global crude oil models. Keeping in mind, that under the assumption that negative production shocks observed in the Iranian production series may not occur, the first shocks are positive and - more importantly - also in absolute terms. The fact that the four-variable SVAR model performed better than the three-variable model in the previous section, when positive flow supply shocks were incorporated into the system is to be stressed again.

Expressing the difference of the counterfactuals in comparison with the observed nominal price of oil, shown in figure 6.9, allows for a better understanding of the dynamics implied in both models as they directly indicate the impact of sanctions had on the nominal price of crude oil. Firstly, comparing the predictions made with the three-variable SVAR model one notes stronger swing movements, as market participants adapted to new conditions. The direction of swing movements is consistent with the theoretical assumptions in chapter 4. Secondly, the implementation of sanctions and also the negative flow supply shocks that were experienced, lead to an increase in the global price of oil of around \$6 per barrel after three months. This decreases global demand for oil and increases production in other parts of the world, which, in turn, lowered prices by \$4 per barrel below the experienced price development after six months. This chain of events continues, but weakens over time, indicating that sanctions have indeed a positive, but temporary effect on the price of crude

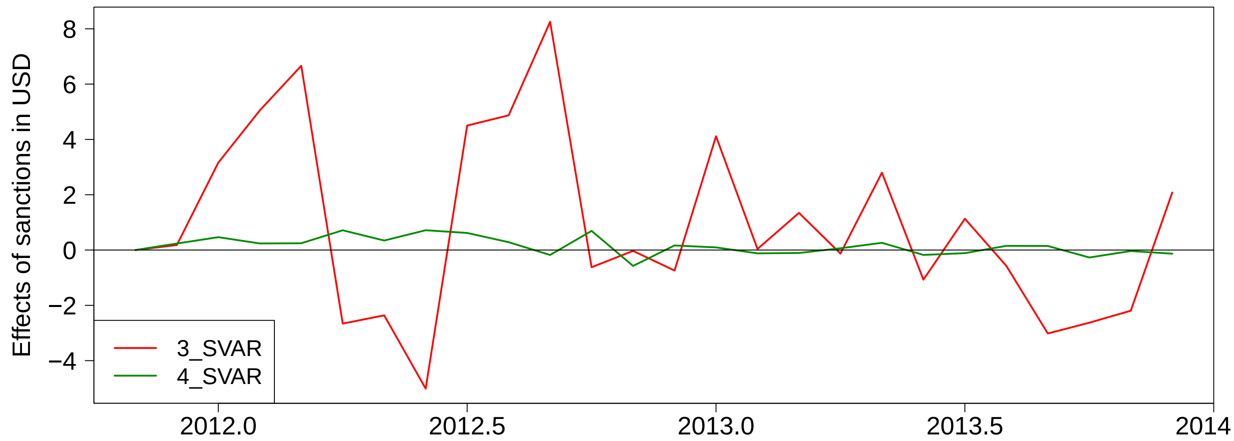

Figure 6.9: Difference between the observed crude oil price and the predicted oil price for both models.

oil. Thirdly, the model estimates that the sanctions induced a higher oil price of around \$8 eight months after their implementation. The four-variable model, however, estimates that the oil price was only marginally increased by sanctions over the evaluation sample. At most it results in an increase of around 40 cents seven months after their imposition.

Although in the previous section the conclusion was drawn that the four-variable model better incorporates the effects of positive flow supply shocks, the estimates achieved with the help of the three-variable model as the swings induced by sanctions correspond more with the expectations of economic theory. More importantly, for both models, the estimates confirm the main arguments in Kilian (2009); Kilian and Murphy (2014) against a major and permanent impact of exogenous flow supply shocks as main determinants of unexpected oil price movements. Even in the case of the three-variable model, all price increases were only temporary and mainly significant within the first half of the evaluation period.

In the next section, the US's recent withdrawal from the JCPOA, sis evaluated to gain insights, whether sanctions may only affect oil prices through flow supply shocks.

6.3.2 Preliminary discussion of the reintroduction of sanctions in late 2018 and the future oil price

Contrary to the previous two sections, here, the attempt is made to qualitatively and quantitatively assess, how the re-imposition of US sanctions after the withdrawal from the JCPOA may affect or may have affected global oil prices. No reliable up-to-date crude oil production data are available so far. However, the attention is drawn to the fact that it is impossible to review this event without taking into account general developments that might additionally affect the global oil price. In the previous sections 4.4.2 and 4.4.3 the historical shock estimates as well as the decomposition of cumulative contributions of each structural shock on the real price of oil were analyzed, with accounting for the fact, that at any given point in time the endogenous variables included in the two models used are affected by all structural shocks simultaneously.

Therefore, it seems, plausible that no significant upward pressure on global oil prices can be expected, if the US withdrawal from the JCPOA induces a corresponding drop in Iranian oil production and exports. Even if the political pressure applied by the US on importers of Iranian crude oil might be stronger than in previous episodes, there exists a natural limit on how much crude oil supplies might be affected by sanctions. In the absence of official statistics on the Iranian oil production during 2019, the EIA estimates that Iranian crude oil exports peaked in June 2018 at around 2.7 mbd.[2] According to oil markets experts, Iran was able to export between 100 and 300 thousand barrels per day in July 2019 (Reuters, 2019a). This implies a substantial and much stronger contraction of oil exports in comparison with the 2011/12 sanctions, that recorded decreases to around 800 thousand barrels per day. These estimates implied a rather negligible influence on the oil price in the case of the four-variable model, and a modest, temporary increase of the price per barrel to $8. Figure 6.10 suggests that the nominal RAP of imported oil fell since July 2018, when Iran's oil exports peaked. Even if the re-imposition of sanctions had an increasing price effect through a negative supply shock (positive effect on the oil price), it was more than compensated by opposing shocks.

The price evolution before the 2011/12 sanctions shows surprisingly the same pattern, prior to the official implementation of sanctions one can observe increasing crude oil prices.

Two possible explanations can be given: Figure 6.10, shows that oil prices were already increasing since January 2016. As shown in figures 4.9 and 4.12 on pages 52 and 56 in chapter 4 the oil price was mainly driven by a combination of flow demand and primarily oil-specific demand (precautionary demand) at the end of the observation period until December 2017. As US President Trump called the JCPOA "the worst deal ever" long before his installation in office in January 2017 (New York Times, 2017), it seems plausible to attribute parts of the oil price increases due to precautionary demand, caused by uncertainties concerning the

[2] https://www.eia.gov/beta/international/analysis.php?iso=IRN (accessed on 3.9.2019).

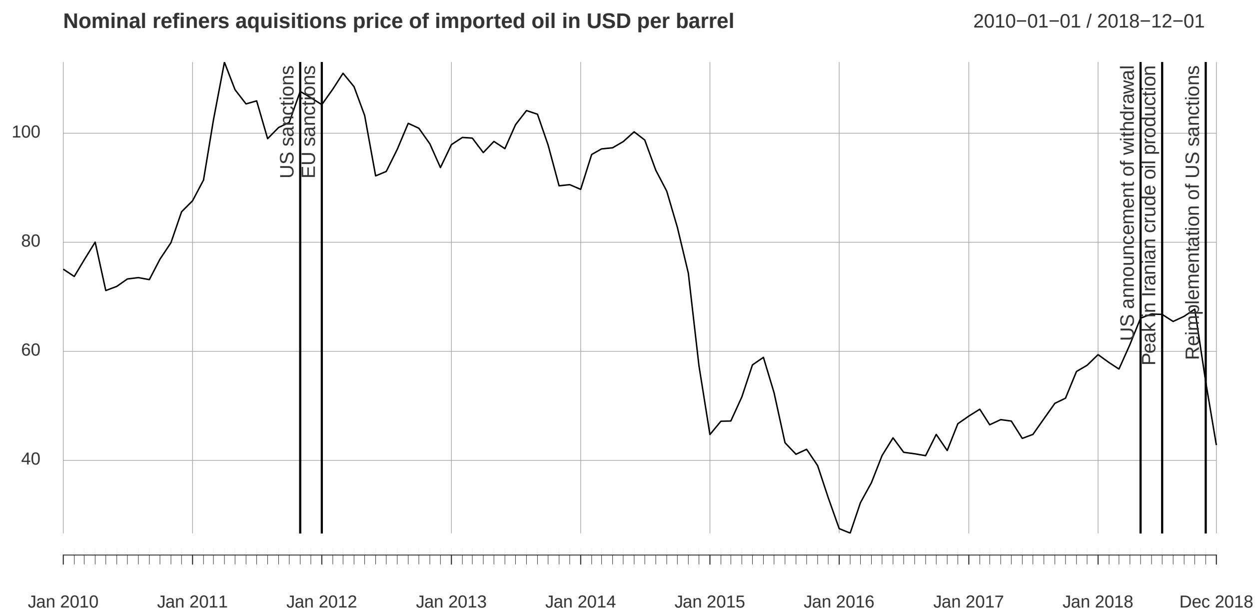

Figure 6.10: Nominal refiners acquisition price of imported oil between January 2010 and December 2018 (Source: EIA).

future Iran policy of the US administration. The election of President Trump resulted in the announcement to withdraw unilaterally from the JCPOA in May 2018. In other words, market participants expected sanctions to be reapplied and reacted accordingly, inducing in an increase in precautionary demand, thus driving crude oil prices upwards. One should note, that precautionary demand might have an even stronger impact on the price of oil, if market participants expect future supply disruptions due to a threatening escalation of the situation in the region.

The order of magnitude of such risks is visualized by the fact, that in 2016 around 18.5 mbd of crude oil passed through the Strait of Hormuz, accounting for around 30% of total global seaborne-traded crude oil (EIA, 2017). So even if a blockade of the strait could be resolved swiftly by military means in case of escalations, oil extracting and oil transportation infrastructure might be critically damaged. The fall of oil prices after the legal imposition of sanctions in 2018 might be a correction of prior precautionary price increases as the market failed to correctly anticipate the effects of supply disruptions.

To investigate this hypothesis we look at the decomposition of unexplained oil price changes between January 2010 and November 2018. We estimate these with the three-variable SVAR framework by extending the estimation sample to November 20018. As shown in figure 6.11, it becomes apparent that the main increases prior to the 2011/12 sanctions as well as the reimplementation of sanctions in 2018 were primarily driven by oil-specific demand shocks. For example, between January 2010 and November 2011 oil-specific demand shocks accounted for a cumulative increase in the real price of oil of around $22 per barrel. On the other hand between January 2017 and June 2018 the cumulative increase due to oil-specific demand shocks in the real price of oil is estimated to around $13 per barrel. We recall that as seen in chapter 4, the main interpretation of oil-specific demand shocks in the three variable-model is primarily precautionary demand because of future developments on the oil markets (Kilian and Murphy, 2014). We thus believe that a part of this identified precautionary demand is due to the Iran policy of the US administration under President Trump.

Given the dynamics within the structural models we think however, that the fall in crude oil prices since July 2018 is more likely the result of a declining business cycle and more importantly negative precautionary demand with regard to future risks associated with the business cycle. The tensions regarding trade policy between the US and China that started in early 2018, as well as the resulting implementation of tariffs and counter-tariffs (see for example Guo et al., 2018; Amiti et al., 2019) directly and indirectly affect the global business cycle. In our models this would be captured by negative flow demand shocks (global demand for industrial commodities including crude oil is affected negatively) as well as negative precautionary demand shocks. As market participants expect the global business cycle to decline, thus expecting crude oil prices to fall, less oil is stored today in order to be consumed tomorrow. These developments however are independent from the Iranian sanctions and need to be viewed on their own. A fall in the cumulative effects of oil-specific demand shocks can be seen in figure 6.11 in accordance with our hypothesis.

In conclusion, we thus believe that the US withdrawal from the JCPOA and the reinforcement of the sanctions had an effect on the price of crude oil primarily through market expectations. Market participants, in anticipation of the reimplementation of sanctions and thus lower Iranian oil exports in the future, were willing to buy and sell crude oil at higher prices before the sanctions took legal effect. As crude oil is easily stored, expectations of future price increases, provide incentives to oil consumers to buy and store oil in the present for future use, thus increasing precautionary demand and by consequence prices today. From an empirical point of view supply shocks are easy to measure and consequently to be used in the SVAR framework. On the other hand, precautionary structural shocks are less easy to historically decompose and need more information.

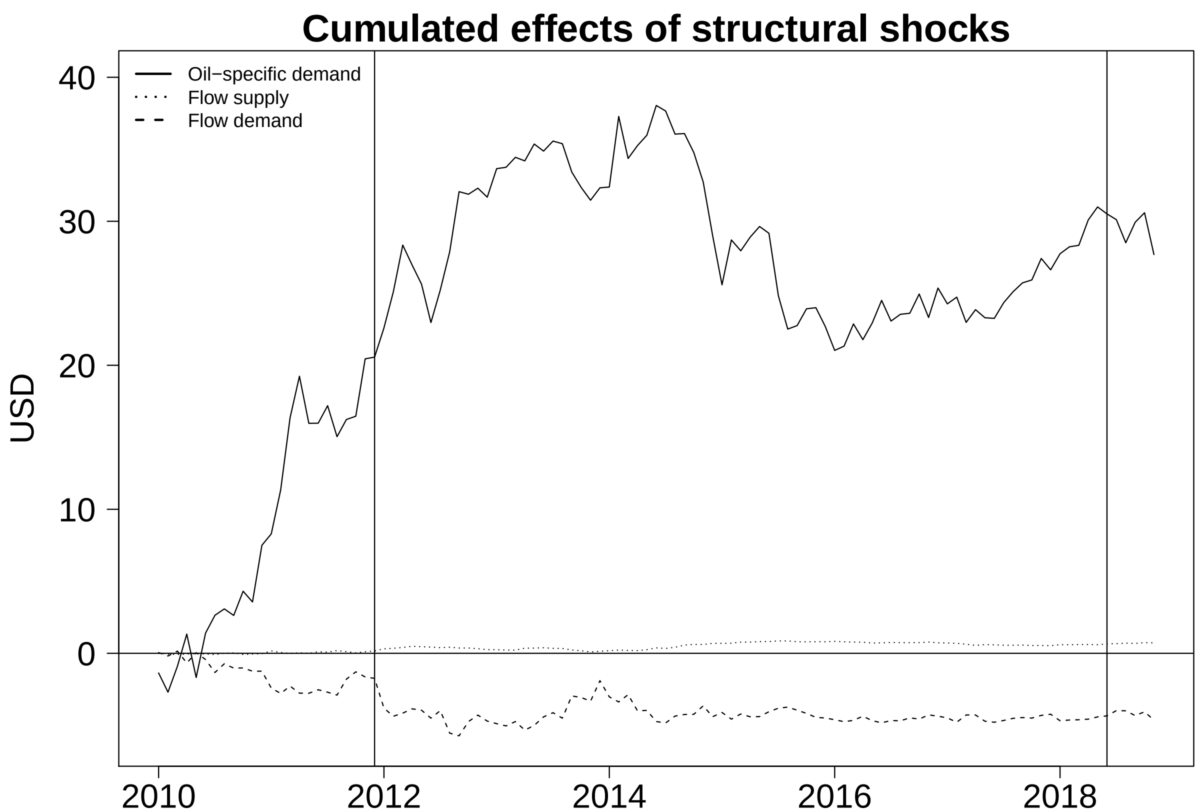

Figure 6.11: Cumulative effects of the estimated structural shocks from the three variable SVAR model between January 2010 and November 2018.

6.4 Conclusion and Discussion

In this chapter, structural global models of oil are used as described in Kilian (2009) and Kilian and Murphy (2014) in order to quantify the effect of an exogenous supply disruption on the global price of oil, induced by sanctions. Work started with the evaluation of forecast performances of the three and four-variable SVAR models in the case of a series of mostly negative shocks, followed by a series of mostly positive oil supply shocks. A first interesting result points at the fact that, when judged by their MSE, no model outperforms the other and vice versa. However, the three-variable model incorporates better information on negative shocks, whereas the four-variable model seems better suited for the incorporation of information on positive shocks.

With the help of both models, two counterfactual evolutions were created for all endogenous model variables in a scenario, in which the negative supply shocks, attributed to the 2011/12 sanctions, never occurred. Similar to a historical decomposition of the observed reduced form shocks, it was possible to isolate the effects of sanctions and estimate an alternative price evolution. This allowed us to quantify the oil price effects of sanctions on Iran by comparing the actual price evolution to both counterfactual price evolutions as predicted by the models. As expected, the preferred model suggested that the sanctions had only insignificant price effects during the course of 25 months. The three-variable model that performed worse in terms of accuracy suggests on the other hand that the sanctions caused the price of crude oil to increase by up to around $8. The effect is however only temporary as other producers quickly reacted and increased their own production to compensate for the Iranian export reductions in the aftermath of the sanctions.

The result suggests that the consequences of sanctions on actors which are not directly involved, but which depend on oil (as importers of foreign oil or exporters of extracted oil) are rather limited, when only supply shocks are taken into consideration.

The comparison and qualitative evaluation of the US withdrawal from the JCPOA and the re-imposition of sanctions in November 2018, however, suggest another dimension through which sanctions might affect the oil price in a more important way. As the intentions of the US administration under President Trump with regard to Iran were and still are rather unclear, it seems not unreasonable to assume that a certain share of the increase in the crude oil price prior to July 2018 was due to an increase in precautionary demand that resulted from the prospect of sanctions on Iran that were soon to come. The same pattern can also be observed for the evolution of prices as well as the cumulative effects of oil-specific demand shock on the real price of oil prior to the 2011/12 sanctions. Thus, while the supply disruptions associated with sanctions appear to be very limited, the effects associated with precautionary demand some time prior to the announcement of formal sanctions seem to be important. Further research in this regard appears promising. Here, we recommend to review further exogenous events associated with measurable disruptions in oil supplies, such as the first and second Gulf Wars in 1991 and 2003, in order to reinforce the argument made in this chapter. If indeed supply disruptions are anticipated by market participants, the structural supply shocks as implied by the global SVAR models might be underrated.

The reference above to the volumes of oil, passing through the Strait of Hormuz in 2016 - about 30% of internationally traded oil - provides another limit to the validation and use of structural models of the three- and four-variable type. All these models are calibrated on the basis of historically observed oil shocks - positive or negative in their implications - and have proven their explanatory power with respect to price oscillations of oil on the world market.

In the framework of renewed sanctions on Iran by the US not only Iran‘s oil and gas production capabilities are heavily affected since 2018, but also increasing hostilities between countries of the region emerged, i.e. between the Schiite Iran and the Sunnite Saudi Arabia. Both are fighting a proxi-war in favour of regional supremacy, actively supported each by one of the military super-powers. The likelihood of an escalation of such hostilities is to be rated high, not the least underfired by recent missile and drone-attacks on Saudia Arabian oil installations. The closure of the Strait of Hormuz - of only 25 km of width - is considered a prime strategic target of Iran in case of an open regional war. Such closure would have to be considered the “Mega-Oil-Shock”, hitherto not known in human history, with not yet imaginable implications on the world oil market. The structural forecast and impact evaluation models do not yet seem well-conceived to handle issues of a regional war in the world‘s most important oil region. A future cooperation and cross-fertilization with structural approaches of international peace research, best represented by the Norwegian Peace Research Institute Oslo and its Founding Director, Johan Galtung, is highly recommended for such purposes.

In the next chapter we will no longer consider the structural models that we relied on to assess historical developments and interdependencies. We will use sparse variable methods in order to evaluate the forecast properties of reduced form VAR models.

Chapter 7

Improving Oil Price Forecasts by Sparse VAR Methods

As already seen, the oil price and its changes have been associated with U.S. macroeconomic aggregates as well as the global business cycle (see e.g. Hamilton (1983), Kilian (2008c)). The important oil price shocks in the 1970s and 1980s gained widespread attention in the public. As of today, crude oil is indispensable for keeping standards of living in developed economies as well as for fueling economic growth in rapidly developing nations such as China and India. Therefore, knowledge about the future price of oil is of importance for different actors. Researchers in central banks and international organizations such as the IMF use oil price forecasts as input in their forward looking macroeconomic models (see Baumeister and Kilian (2014b)). Thus, improving crude oil price forecasts helps generating better macroeconomic projections as well as better future risk assessment associated with oil price fluctuations. Oil price forecasts are also helpful for governments of oil exporting countries which strongly depend on oil revenues to finance public expenses in budget planning. Relatedly, forecasting the oil price aids governments of countries that heavily rely on crude oil imports in shaping their environmental policies and energy tax setting. Improved oil price forecasts also support firms in their investment and purchasing decisions. For example, airlines and automobile companies take oil price forecasts into consideration when they decide about fares and product prices as well as product portfolios. Similarly, private homeowners might upgrade to energy-saving heating systems when forecasts point to future heating oil price increases.

In this chapter we start from the global three-variable VAR model for crude oil as first proposed by Kilian (2009) and presented in chapter 4 as the benchmark and investigate variants

This chapter is based on joint work with Jens Krüger.

with enhanced variable sets using sparse (regularization) methods in order to evaluate and compare their forecast properties. Sparse estimation methods gained widespread attention in the machine learning literature (see e.g. Murphy, 2012) and now find more and more economic applications as a variable selection procedure. This is particularly important for VARs where a large number of parameters are to be estimated and usually only a common lag length for all equations is selected by information criteria.

Also in previous research on oil price forecasting an increasing trend towards basing the forecasts on a broader information set can be observed. This strand of research is mostly focused on applications of forecast combination methods see (see Baumeister et al., 2014; Baumeister and Kilian, 2015; Funk, 2018; Garratt et al., 2019; Wang et al., 2017; Zhang et al., 2018). More recently neural networks as well as regularization methods also have been employed to improve oil price forecasts (see Cheng et al., 2019; Zhang et al., 2019).

The lessons from the forecast evaluation exercise reported in this chapter can be summarized as follows. First, the results show that the lag order commonly fixed at 12 or 24 months, which is as mentioned in chapter 4 justified for impulse-response analysis, is detrimental to forecast performance. Second, appropriate variable transformation (logs, differences, levels) is crucial for the forecast performance. Third, applying sparse estimators leads to improvements in forecast performance when using the variable transformation originally employed by Kilian (2009) and in the VAR in levels. Regularization also improves forecasts for shorter horizons when we express the variables in differences. Finally, when augmenting the core variable set by industrial production indices, exchange rates and financial variables, regularization does not lead to forecast improvements and we even observe forecast deterioration in some occasions.

The chapter unfolds as follows. Section 2 introduces the VAR framework as well as the three sparse estimation methods which are subsequently applied. Section 3 presents the core data series and discusses data transformation and stationarity assessment. In section 4 we discuss different basic VAR specifications and select the best performing one as the benchmark. We proceed by estimating the three-variable VAR with the sparse methods and evaluate the forecast performance in comparison to the selected benchmark. Section 5 extends the three core variables by further variable sets containing production indices, exchange rates, financial variables and impulse indicator saturation dummies and evaluates the forecast performance. We conclude in section 6 with the discussion of the main findings.

7.1 Sparse VAR Methods

Before we turn to the description of the data used and how we dealt with stationarity-integration issues we briefly explain the forecasting models used. Our forecast evaluation exercise relies on the framework of a reduced form VAR as presented in chapter 4. In addition, we explain the approaches to regularization which we employ to prune the parameter matrices to obtain a more parsimonious (sparse) model specification, more precise parameter estimates and possibly reduced forecast errors. It is the primary aim of this study to investigate the latter issue.

Let's recall a VAR stated as a VAR(p) with p lags for m variables in the vector $\boldsymbol{y}_t = (y_{1t}, ..., y_{mt})'$ observed for the periods $t = 1, ..., T$,

$$\boldsymbol{y}_t = \boldsymbol{c} + \boldsymbol{A}_1 \boldsymbol{y}_{t-1} + ... + \boldsymbol{A}_p \boldsymbol{y}_{t-p} + \boldsymbol{u}_t. \tag{7.1}$$

A VAR can be consistently estimated by least squares equation by equation, which amounts to minimizing the sum of squared residuals

$$SSR(\boldsymbol{\theta}) = \sum_{t=p+1}^{T} \boldsymbol{u}_t' \boldsymbol{u}_t \tag{7.2}$$

as the objective function, where the parameter vector $\boldsymbol{\theta}$ is understood to stack all $k = m + pm^2$ parameters to be estimated (i.e. $\boldsymbol{c}$ and $\boldsymbol{A}_1, ..., \boldsymbol{A}_p$).

Given the estimates for $\boldsymbol{c}$ and $\boldsymbol{A}_1, ..., \boldsymbol{A}_p$, denoted $\hat{\boldsymbol{c}}$ and $\hat{\boldsymbol{A}}_1, ..., \hat{\boldsymbol{A}}_p$, respectively, the VAR can be used for generating forecasts by iterating equation (7.1) forward. This leads to forecasts one step and two steps into the future, written as

$$\hat{\boldsymbol{y}}_{T+1|T} = \hat{\boldsymbol{c}} + \hat{\boldsymbol{A}}_1 \boldsymbol{y}_T + ... + \hat{\boldsymbol{A}}_p \boldsymbol{y}_{T-p+1} \tag{7.3}$$

and

$$\hat{\boldsymbol{y}}_{T+2|T} = \hat{\boldsymbol{c}} + \hat{\boldsymbol{A}}_1 \hat{\boldsymbol{y}}_{T+1|T} + ... + \hat{\boldsymbol{A}}_p \boldsymbol{y}_{T-p+2}, \tag{7.4}$$

respectively. Here $\hat{\boldsymbol{y}}_{T+1|T}$ denotes the forecast for the variables one time step into the future given that the available information ends in period T. Note that for the 2-step forecast $\hat{\boldsymbol{y}}_{T+2|T}$ the first lag on the right hand side would be $\boldsymbol{y}_{T+1}$ which is not available in the data (the sample ends in period T) and is therefore substituted by the 1-step forecast $\hat{\boldsymbol{y}}_{T+1|T}$. In general, the h-step forecasts generated by conditional expectations are estimates of the conditional expectation $E(\boldsymbol{y}_{T+h} \,|\, \boldsymbol{y}_T, ..., \boldsymbol{y}_1)$. The h-step forecasts are computed by

$$\hat{\boldsymbol{y}}_{T+h|T} = \hat{\boldsymbol{c}} + \hat{\boldsymbol{A}}_1 \hat{\boldsymbol{y}}_{T+h-1|T} + ... + \hat{\boldsymbol{A}}_p \hat{\boldsymbol{y}}_{T+h-p|T}, \tag{7.5}$$

upon the substitution $\hat{\boldsymbol{y}}_{T+j|T} = \boldsymbol{y}_{T+j}$ whenever $j \leq 0$. Forecasts constructed in this way minimize the theoretical MSE.

The number of parameters arising in unconstrained VAR with lag length p is usually quite large, i.e. $k = m + pm^2$. Not all those parameters are different from zero although their estimates are so by chance and this may be detrimental to forecast performance. Since information criteria for lag order selection only eliminate entire parameter matrices $\boldsymbol{A}_j$, it would be helpful to use statistical methods which constrain selective parameters within these matrices to be zero. In the statistical literature this is known as sparsity or regularization to reduce the number of parameters which are different from zero.

Typical methods for regularization are the least absolute shrinkage and selection operator (LASSO), the Elastic Net (ENET) and the smoothly clipped absolute deviations (SCAD) method which are explained below. These methods have in common that a penalty term $P(\boldsymbol{\theta})$ for the magnitude of the parameters is added to the objective function to be minimized

$$Z(\boldsymbol{\theta}) = SSR(\boldsymbol{\theta}) + \lambda P(\boldsymbol{\theta}) \tag{7.6}$$

with the penalty weight $\lambda > 0$ to be determined by cross-validation techniques. Hastie et al. (2009) provide a lucid exposition of variable selection by regularization methods (also known as shrinkage methods) in general.

In this work we investigate the forecast performance of VARs estimated by three common variants of regularization methods. First, the LASSO by Tibshirani (1996) specifies the penalty term as $P(\boldsymbol{\theta}) = \sum_{j=1}^{k} |\theta_j|$.[1] This constrains some of the parameter estimates to be exactly equal to zero and thus eliminates some of the lags of the corresponding variables in the VAR to reach sparsity. Second, the ENET by Zou and Hastie (2005) chooses $P(\boldsymbol{\theta}) = \sum_{j=1}^{k} \left(\alpha|\theta_j| + (1-\alpha)\theta_j^2\right)$ which is the a combination of the LASSO and Ridge penalties with α usually fixed at 0.5. Third, SCAD by Fan and Li (2001) is based on $P(\boldsymbol{\theta}) = \sum_{j=1}^{k} p(\theta_j)$ with $p(\theta_j) = |\theta_j|$ if $|\theta_j| \leq \lambda$, $p(\theta_j) = (2\gamma|\theta_j| - \theta_j^2/\lambda - \lambda)/2(\gamma - 1)$ if $\lambda < |\theta_j| \leq \gamma\lambda$ and $p(\theta_j) = \lambda(\gamma + 1)/2$ if $|\theta_j| > \gamma\lambda$ with $\gamma > 2$ (setting $\gamma = 3.7$ is recommended by (Fan and Li, 2001, p. 1351) as providing "good practical performance for various variable selection problems"). The SCAD penalty coincides with the LASSO for $|\theta_j| \leq \lambda$, is a concave quadratic function until $|\theta_j| \leq \gamma\lambda$ and is constant for $|\theta_j| > \gamma\lambda$. This relaxes the intensity of penalization

[1] This is in contrast to Ridge regression, introduced by and Hoerl and Kennard (1970), using the penalty term $P(\boldsymbol{\theta}) = \sum_{j=1}^{k} \theta_j^2$ which serves to shrink the parameter estimates towards zero but does not set some of them exactly equal to zero as the LASSO does.

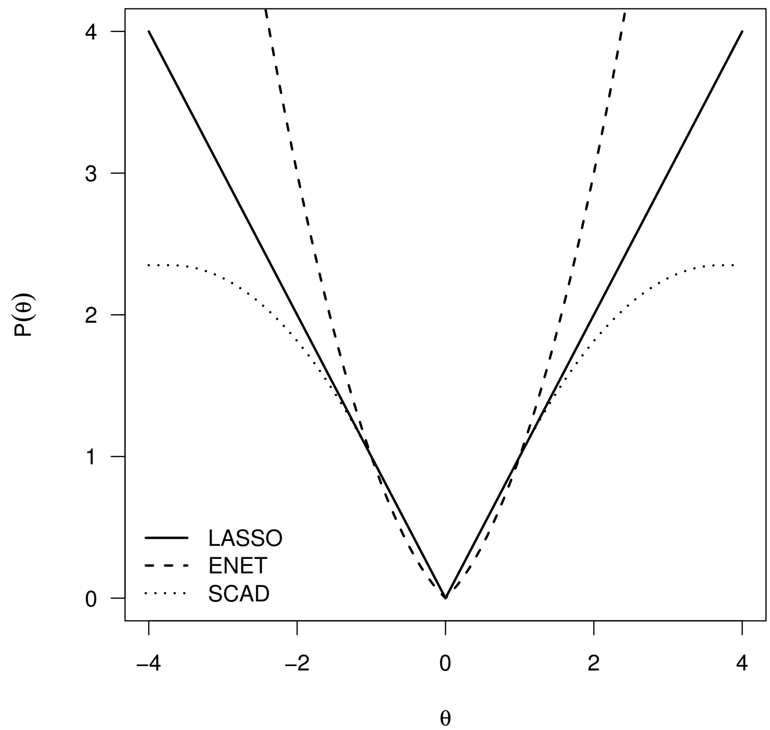

Figure 7.1: Penalty Functions.

when the absolute value of the parameter increases.[2]

Figure 7.1 shows the penalty functions for the three sparse variants considered depicted for a scalar parameter θ (setting $\lambda = 1$, $\alpha = 0.5$ and $\gamma = 3.7$). All computations in this chapter are performed using the packages "vars" and "sparsevar" for R.

7.2 Data and Stationarity

In this section we briefly review and discuss of the three core variables used in the VAR models in section 7.3 and previously presented in chapter 4. As already seen, these variables are used by Kilian (2009) in his structural VAR model to capture the main dynamics of the global market for crude oil as well as to estimate historical oil price shocks. The main difference in contrast to chapters 4-6 is that we use different transformations and evaluate the resulting forecasting properties. The same original variables are also used for the evaluation of oil price forecasting in the handbook article by Alquist et al. (2013). Later, in section 7.4,

[2] A quite similar suggestion, called minimax concave penalty (MCP), has been made by Zhang (2010), which leads to results which are almost indistinguishable from SCAD and is therefore not further considered in the empirical forecast evaluation exercise of this chapter.

these variables are extended by further sets of variables.

The real oil price is one of the three core variables used by Kilian (2009) and Alquist et al. (2013). More specifically, the three core variables are the real price of crude oil (deflated by the US consumer price index and expressed in logs) rpo_t, the index of global real economic activity as by Kilian (2009) rea_t and the percentage change in global crude oil production (computed as log differences) $\Delta prod_t$. See Kilian (2009) for a more thorough discussion of the construction of the variables, in particular regarding the real activity index. The index is constructed using dry cargo shipping rates based on the idea that global economic activity is the main driver of demand for international freight transport services. The updated data are retrieved from the homepage of Lutz Kilian (http://www-personal.umich.edu/~lkilian/) and incorporate the updates addressed in Kilian (2018) based on the methodological critique of Hamilton (2018).

The time series of the three core variables are plotted in figure 7.2. The sample period in this chapter spans January 1974 to December 2017 implying a total sample size of $T = 528$ months. This extends the sample period of Kilian (2009), which goes until December 2007, now also comprising the time of the financial crisis, the breakdown of Lehman Brothers, the Great Recession and the recovery thereafter. For the real price of crude oil we use the refiners acquisition price of imported oil deflated by US CPI proposed by Kilian (2009) as the best measure for global oil prices.[3]

In the first panel the real oil price is expressed in logs. The first row of the figure shows a trend in the production series and long swings of the real oil price and to a lesser degree in the case of the real activity index, pointing to a substantial degree of persistence. Both trend and persistence are characteristics of unit root nonstationarity (with and without a drift component, respectively). Therefore, this visual inspection suggests using the transformations (Δrpo_t, Δrea_t, $\Delta prod_t$) for the three core variables in the VAR.

When we try to confirm this by formal statistical testing the results (not shown in detail here) are mixed. For all three variables we find a strong rejection of the stationarity null hypothesis using the KPSS test of Kwiatkowski et al. (1992). This test is, however, prone to severe size distortions and therefore leads to substantial overrejections of the null hypothesis also under stationarity. The unit root null hypothesis is, however, also rejected in the case

[3]The refiners acquisition price of imported oil and global oil production series (in thousand barrels per day) are retrieved from the US Energy Information Administration. US CPI is retrieved from the FRED database with the series code CPIAUCSL.

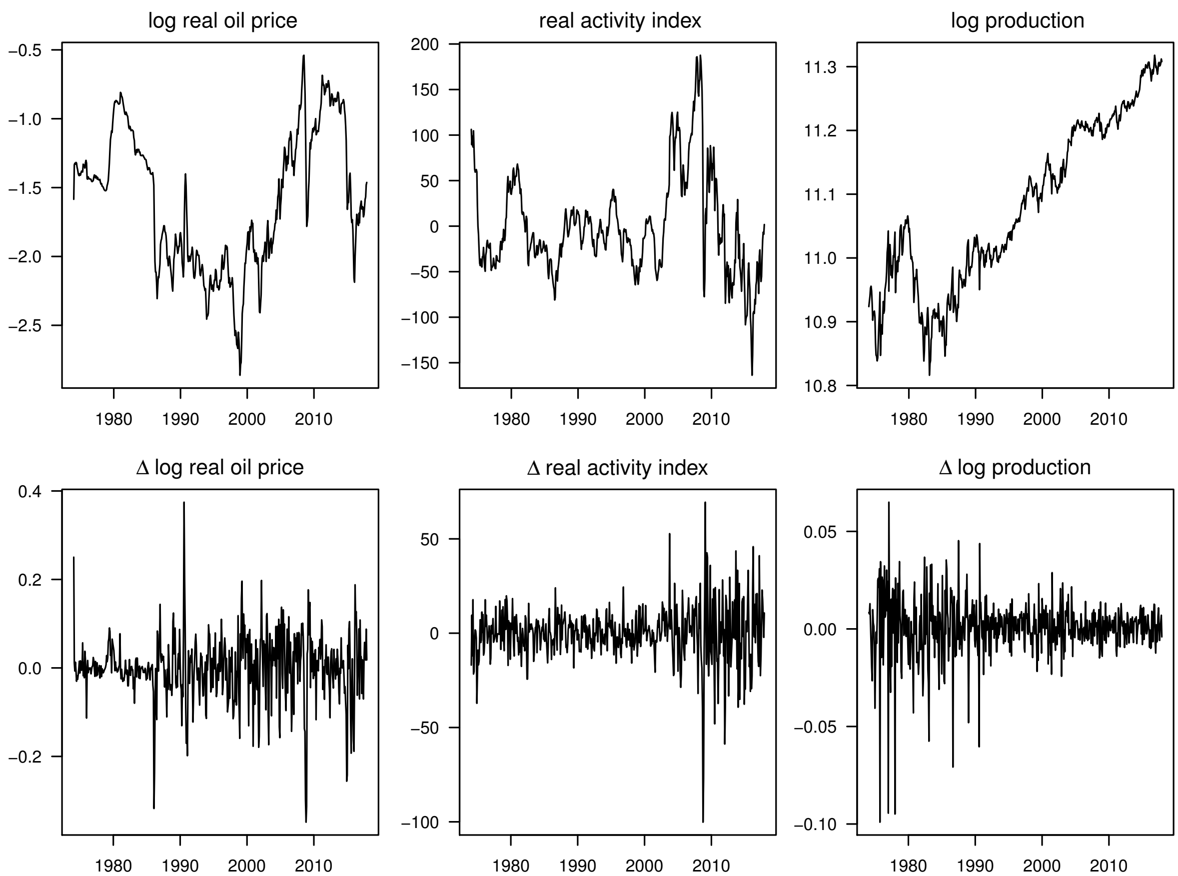

Figure 7.2: Time series of the Core Variables.

of the (log) real oil price using the DF-GLS test (or ERS test) of Elliott et al. (1996). For the real activity index and (log) oil production, the unit root null can not be rejected. This is not overly surprising for the production series, but is somewhat puzzling in the case of the real activity index and its appearance in the figure 7.2. Applying the testing procedure to the first differences of the three variables we observe strong rejections of the unit root null, jointly with no rejections of the stationarity null. This is consistent with the visual inspection.

For control purposes and to achieve consistency with the literature (especially Kilian (2009) and Alquist et al. (2013) we also perform the forecast evaluation exercise for the transformations (rpo_t, rea_t, $\Delta prod_t$), where, as defined above, the real price of oil and the production are expressed in logs. Regardless of the transformations applied, the target variable of the forecasts is the real price of oil (unlogged) which is the variable decisions makers are most likely to focus on rather than the corresponding logs or growth rates (differences of logs). According to Sims et al. (1990) determining the correct order of integration is not problematic for consistent parameter estimation in VAR models and should therefore not be problematic for forecasting.

7.3 Results with the Core Variables

In this section we discuss the results from the forecast evaluation exercise based on the three core variables. We first compare VARs with lag lengths fixed at $p = 12$ and $p = 24$, a VAR with lag length selected by Akaike information criterion (AIC), and naïve no-change prediction (average real oil price over the previous 12 months) to select a benchmark for the sparse VAR methods. In the second step we evaluate the performance of the sparse VARs, estimated by LASSO, ENET and SCAD, in comparison with this benchmark model for the forecast horizons $h = 1, 2, 3, 6, 9, 12$.

Regarding the transformations of the three core variables we distinguish the Kilian VAR with variables $\boldsymbol{y}_t = (rpo_t,\, rea_t,\, \Delta prod_t)'$ as analyzed in Kilian (2009) and Alquist et al. (2013), the VAR in differences with variables $\boldsymbol{y}_t = (\Delta rpo_t,\, \Delta rea_t,\, \Delta prod_t)'$ and the VAR in levels with $\boldsymbol{y}_t = (rpo_t,\, rea_t,\, prod_t)'$. Using a VAR in levels, irrespective of the orders of integration of the variables and possible cointegration among the variables, is a standard approach in some fields, e.g. the empirical assessment of monetary policy (see e.g. Christiano et al., 2005). There is a also considerable literature on the forecast performance of VARs in levels versus first differences (see e.g. Hoffman and Rasche, 1996).

When we suppose that all variables are integrated of order one and we are indeed able to establish cointegration by the Johansen (1988, 1991) trace test. Using an expanding test sample size starting from the first 100 observations up to the total sample we can establish cointegration for most of the samples before the financial crisis which is substantially weakened by the impact of the crisis. In the presence of cointegration the Granger representation theorem (Engle and Granger, 1987) justifies the estimation of a VAR in levels as a reduced form basis for forecasting. Even in the absence of cointegration there are good arguments that the decision between differences or levels is rather inessential when the VAR is used for forecasting. As explained by Kilian and Lütkepohl (2017, pp. 373f.) the main reason is the inherent ability of the VAR in levels to encompass a VAR model with integrated and possibly cointegrated variables as well as a VAR for stationary time series. This argument is reinforced by the uncertainty about unit root and cointegration properties of the time series and the often neglected fact that deciding between a VAR in differences and a cointegrating VAR is also subject to pre-testing bias.

Depending on the specific transformation of the real oil price, we obtain a forecast of the log (Kilian VAR and VAR in levels) or of the log differences (VAR in differences) of the real oil price variable. To compare these forecasts with the unlogged real oil price as our target variable, the forecasts are appropriately re-transformed (meaning taking exponentials when the real oil price has been logged or cumulating growth rates starting from the last observation in the data).

The forecast experiment is specified with an expanding window for the estimation sample with the first sample spanning 20 years (240 months) from January 1974 until December 1993 and the first forecast for January 1994 for a horizon $h = 1$ (February 1994 for $h = 2$, March 1994 for $h = 3$, June 1994 for $h = 6$, September 1994 for $h = 9$ and December 1994 for $h = 12$). Note that in the subsequent figures all forecast error measures are aligned at the position of the final observation of the estimation sample (i.e. December 1993 in the case of the first forecast) irrespective of the forecast horizon. Then the procedure is repeated with a further month, January 1994, added to the estimation sample. Proceeding in this way month by month we end up with a final estimation sample from January 1974 until December 2016 (43 years or 516 months) with forecasts for January 2017 ($h = 1$) until December 2017 ($h = 12$) which are all assigned to December 2016 in the figures.[4]

[4] Some corresponding results with a rolling window design of the forecast experiment (in fact a rolling window of 240 months) are collected in the appendix.

7.3.1 Benchmark VAR

The results for four candidates of our benchmark model are shown in figure 7.3. The curves show the recursive MSE measures[5] for the VAR(24) with a fixed lag length of $p = 24$ (VAR24, dotted line), used by Alquist et al. (2013), the VAR(12) with reduced lag length of $p = 12$ (VAR12, dashed line), VAR(AIC) with the lag length p chosen by the Akaike Information Criterion (VARAIC, dash-dotted line)[6] and the naïve no-change forecasts (solid line), which are used as the benchmark forecast in Alquist et al. (2013).

Each column pertains to a different transformation of the three variables (from left to right: VAR with transformation as in Kilian (2009), VAR in differences, VAR in levels) while the rows show the results for a particular forecast horizon of $h \in \{1, 2, 3, 6, 9, 12\}$ months. The horizontal lines indicate the smallest recursive MSE value at the end of the evaluation period which is achieved by any of the methods under consideration. The numerical value of this smallest MSE is printed directly above the horizontal line.

What we observe at first is the general tendency of a steady increase of the MSE over time. Thus the accuracy of the oil price forecasts deteriorates systematically since the 1990s. This might be explained by the several changes affecting global oil markets in the late 1980s (see Hamilton, 2009b). The collapse of OPEC had lasting implications. The powerful cartel from the 1970s never recovered from the oil price collapse in 1986 and permanently lost influence on global markets. The fall of the Soviet Union and the emergence newly independent oil producing countries was a further source of oil market disruptions. The second obvious characteristic is the impact of the financial and economic crisis with the consequence of a series of particularly bad forecasts, leading to a pronounced rise of the MSE lines. After about 2010 forecast errors stabilize on a high level or appear to to improve by a small margin.

The central column of the figure clearly shows that MSE values obtained with a VAR in differences are generally smaller than those obtained with the Kilian VAR and the VAR in levels across all specifications. However, looking down the columns of the figure we observe that the forecast performance quickly deteriorates with increasing forecast horizon. For the VAR in differences the VARAIC is the best forecasting method, closely followed by VAR12

[5]Depicted is $MSE_t = t^{-1}\sum_{s=1}^{t}(y_s - \hat{y}_{s,h})^2$ with the realization of the real oil price denoted by y_s (not in logs) and the h-step forecast $\hat{y}_{s,h}$ for the same period s, obtained by a particular method (indicated in the legend in the first row of the figure) and appropriately retransformed from of the variables included in the VAR.

[6]Using the Bayesian information criterion (BIC) leads to very similar lag length selection and very similar results.

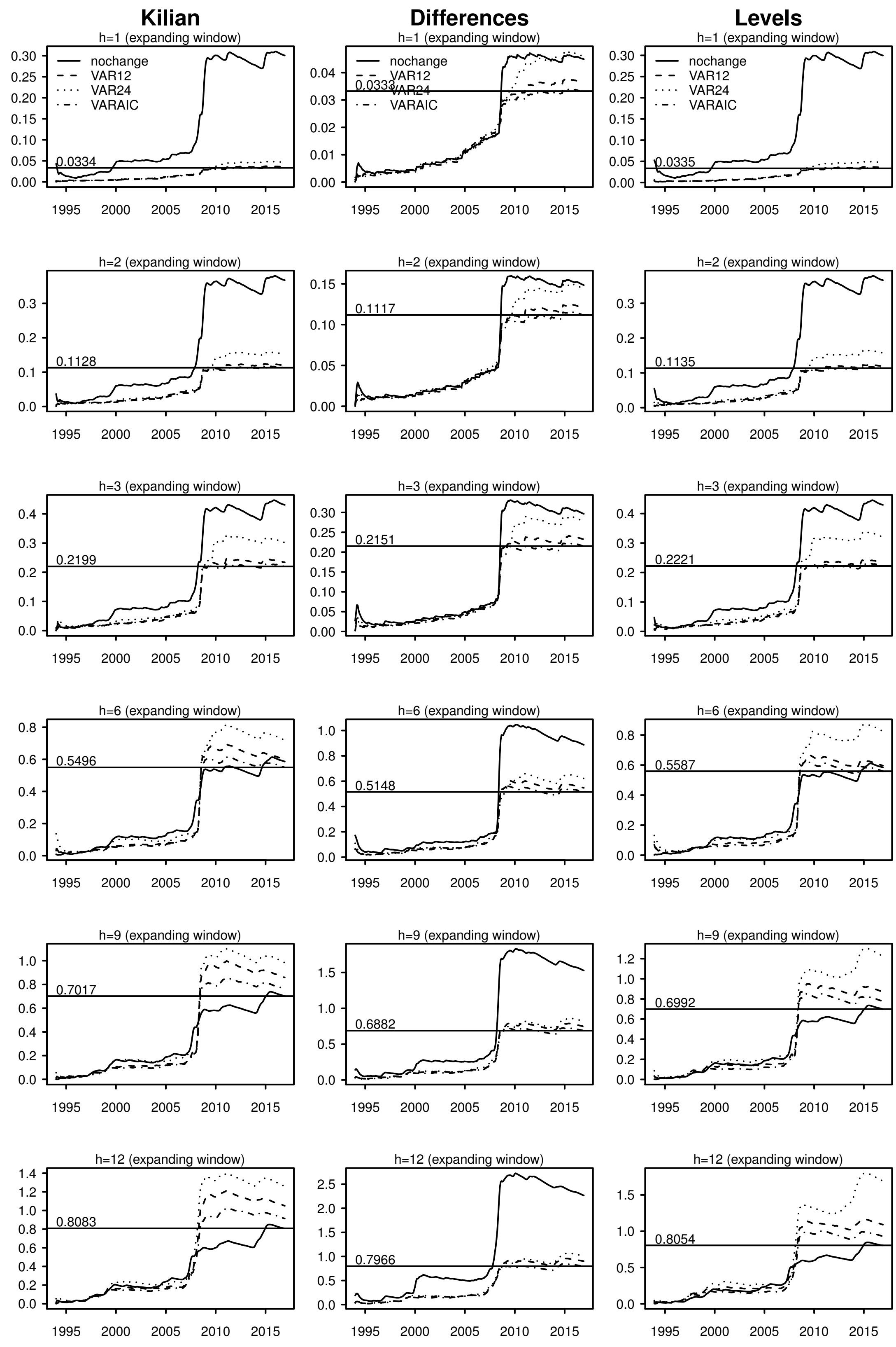

Figure 7.3: Benchmark selection (expanding window).

and VAR24. The left and right columns in the figure, pertaining to the Kilian VAR and the VAR in levels, respectively, roughly contain the same message. For the shorter forecast horizons ($h = 1, 2, 3$) the VARAIC performs better than the VARs with a lag order fixed at 12 or 24, while the no-change forecast performs worst. In contrast, for the longer forecast horizons ($h = 6, 9, 12$), the VAR12 and VAR24 perform poorly, while there is a close competition of the VARAIC and the naïve no-change forecasts with the no-change forecasts becoming slightly better at the longest forecast horizons.

Taking these results together we select the VARAIC as the overall best forecasting method and decide to use this method as the benchmark in the subsequent comparison with the sparse VAR approaches.[7] This allows for a direct comparison of the effects of the regularization (imposing sparsity) within the common framework of a VAR model. The main issue is the distinction of pruning entire coefficient matrices versus pruning single coefficients within these matrices.

7.3.2 Sparse VARs

Figure 7.4 is an analogous depiction of the results for the sparse VAR models, i.e. the basic LASSO, ENET and SCAD, shown by the solid, dashed and dotted black lines, respectively. The recursive MSE of the benchmark VARAIC is shown as gray lines. We start with a VAR(12) to which the regularization is applied. Note that the regularization in equation (7.6) depends on the relative magnitudes of the parameters which in turn depends on the scaling of the variables. Thus, all variables are normalized to have the same standard deviation, which is the standard deviation of the log real oil price.

As before, we find the same general increase of the forecast error measures, especially during the months of the financial crisis. Skimming through the forecast horizons in search for the best combination of variable transformation and estimation method we first of all observe that the forecasts from a VAR in differences remain slightly better than those obtained from the VAR with variables transformed according to Kilian and the VAR in levels at shorter forecast horizons but loses ground at longer forecast horizons. This holds irrespective of the particular form of regularization used for the VAR estimation.
Comparing the Kilian VAR and the VAR in levels we find that the MSE at the end of the

[7] We also estimated vector error correction models (VECM) with imposing cointegration relations determined by the Johansen (1988, 1991) methodology. The forecast errors do not point to an improvement of the predictive performance compared to the VARAIC.

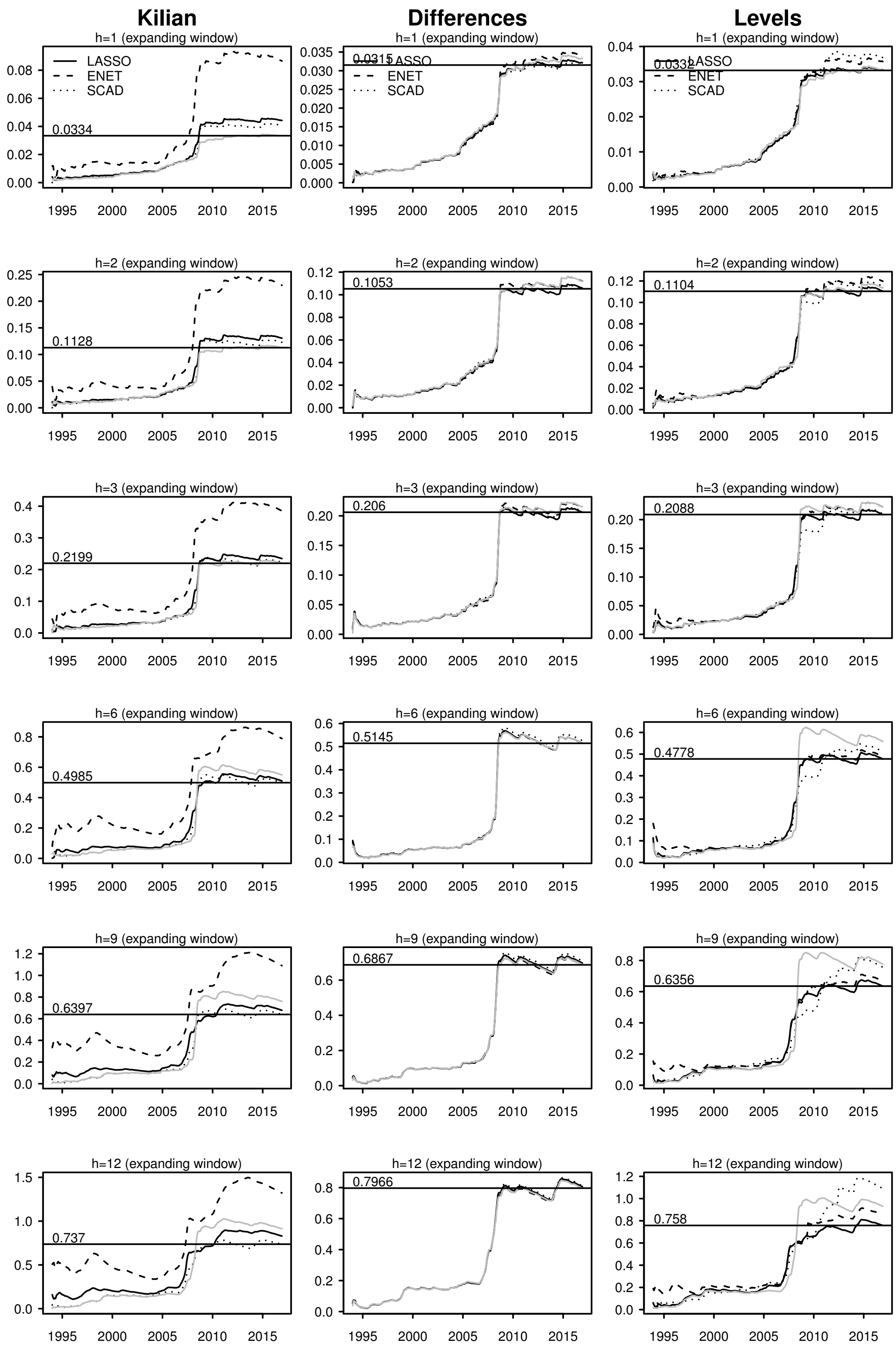

Figure 7.4: Evaluation of the sparse VARs (expanding window).

forecast period obtained with the best method (the number above the horizontal line) is smaller in the case of the VAR in levels for all forecast horizons. The particular estimation method which reaches the smallest MSE at the end differs, however. For the shorter forecast horizons ($h = 1, 2, 3$) and the Kilian VAR the VARAIC is best, closely followed by SCAD and LASSO, while ENET performs worst. In the case of the VAR in levels the ranking is different. Here, the LASSO and the ENET are the best methods and are close to each other. The VARAIC is also close for the shortest forecast horizons with a widening gap when the forecast horizon increases. The SCAD method performs worst for all forecast horizons.

For the longer forecast horizons ($h = 6, 9, 12$) the LASSO and ENET are best in the VAR in levels and overall. For the VAR in levels the VARAIC and SCAD get worse with increasing forecast horizon. Interestingly, VARAIC and SCAD perform best before the financial crisis and also perform better than the other methods when the forecast performance worsens during the period of the financial crisis. SCAD remains best until about 2010 as is visible by the dotted line being lower than the other lines. The ENET is the worst method before the financial crisis and gains much in performance afterwards. Again, there are differences when the Kilian VAR is considered. With this transformation, SCAD and LASSO are best at the end of the sample period and also perform quite well before, in particular since the financial crisis. While VARAIC and SCAD are best performing before the financial crisis, VARAIC loses much more performance during the financial crisis than SCAD does. The ENET generally performs worst in the Kilian VAR.

Considering all results together, we see that the forecast performance depends on the transformation of the variables included in the VAR and there are also pronounced differences across all employed regularization approaches to induce sparsity. Most important, if a particular regularization method performs well with a particular transformation of the variables this does not imply that the same method also performs well with a different variable transformation. The comparison of ENET and SCAD shows this clearly. The basic LASSO appears to be a quite good allrounder which not always performs best but adapts well to different transformations of the variables and is never far behind the best performing method.

7.4 Extended Variable Sets

One of the main virtues of sparse regression methods is the property to deal with situations in which there are more variables than observations. This is enabled by the least-angle regression (LARS) algorithm (Efron et al., 2004) which can cope with those situations (see Hastie et al., 2009, ch. 18, for an exposition). When we extend the variable set consisting of the three core variables considered so far by further variables, the number of parameters grows with the square of the number of variables in the VAR for a constant lag length. Thus, it is of particular interest to investigate whether the sparse VAR methods are able to exploit the predictive power of further variables in extended variable sets.

In this section the VAR model with the three core variables is extended by different variable sets containing industrial production indices of the G7 countries[8], exchange rates to the US dollar and variables related to different investment opportunities. Further more, we also apply impulse-indicator saturation (IIS) to eliminate the potentially adverse effects of single observations on the forecast performance.

The panels in subsequent figures are arranged analogous to the previous section. Now the black lines show the cumulative MSE values with the extended variable sets (solid for the LASSO, dashed for ENET and dotted for SCAD). The gray lines represent the corresponding results only including the three core variables as discussed above (see figure 7.4) for the purpose of a direct comparison of the effects of the enlarged variable sets.

7.4.1 Production Indices

The theoretical reasoning behind these additional variables is obvious in the case of the industrial production (or changes thereof) which is a major driver of the oil price. Admittedly, large newly industrializing countries like China or India are not included for reasons of data availability. Although this omission is not critical at the start of the sample period it may become increasingly crucial nearing the end of the sample period. As far as the industrial production in these countries is linked to the industrial production of the G7 countries this omission can be accommodated by the VAR coefficients. The data for industrial production are retrieved from the FRED database.[9] Under our three transformations of the variables we take log differences of the production indices in the case of the Kilian VAR and the VAR in differences, whereas we use log levels in the case of the VAR in levels.

Adding the production indices of the G7 countries leads to the results shown in figure 7.5 under an expanding window design. Considering first the Kilian VAR with the growth rates

[8] Canada, France, Germany, Italy, Japan, United Kingdom and United States.

[9] The respective codes are CANPROINDMISMEI, FRAPROINDMISMEI, DEUPROINDMISMEI, ITAPROINDMISMEI, JPNPROINDMISMEI, GBRPROINDMISMEI and INDPRO.

(computed as log differences) of the production indices added we observe that the smallest final MSE values are reached by the SCAD method for all forecast horizons considered. The results are almost indistinguishable from the previous results without including the production indices (in the figure the gray dots for the SCAD results are almost completely plotted over the black dots from the extended model). It seems that the additional variables are completely pruned out by SCAD regularization. The other regularization methods, i.e. LASSO and ENET, are associated with larger final MSE values when the production indices are included.

In the case of the VAR in differences we observe no smaller MSE across all forecast horizons compared to the results of the previous section. We also find substantially larger MSE values across all horizons when the VAR in levels is extended by the log levels of the production indices. Here, the increase in MSE is so large that the ranking of the Kilian VAR and the VAR in levels reverses. Specifically, the Kilian VAR is now better than the VAR in levels in terms of final MSE (black lines) but is not better than the VAR in levels without the extension by the production indices (gray lines).

7.4.2 Exchange Rates

In parallel with the previous subsection we now extend the variable set by the exchange rates of the G7 countries (excluding the US) to the US$. Since the international oil trade is conducted in US$, it makes sense to extend the VAR by the exchange rates of the G7 countries (excluding the US) to the US$. This is especially true when we consider that most G7 countries heavily rely on imported oil traded in US$ to satisfy domestic demand. The theoretical justification for adding the exchange rates is based on models such as Krugman (1983a,b). For empirical work on the relationship between crude oil prices and real exchange rates we refer to Zhou (1995); Amano and van Norden (1995); Benassy-Quere et al. (2007). The series on the exchange rates are also taken from the FRED database.[10] We apply the same transformations as in the case of the production indices in the previous subsection.

The results are reported in figure 7.6. The extension of the Kilian VAR by the log differences of the exchange rates gives rise to similar results as the extension by industrial production

[10] The respective codes are EXCAUS, CCUSMA02FRM618N, CCUSMA02DEM618N, CCUSMA02ITM618N, CCUSMA02GBM618N and EXJPUS. For France, Germany and Italy we point out, that exchange rate is expressed in Euro to US$. Before the introduction of the common currency in January first, 1999 the series are constructed by using the official national fixed exchange rates to the Euro.

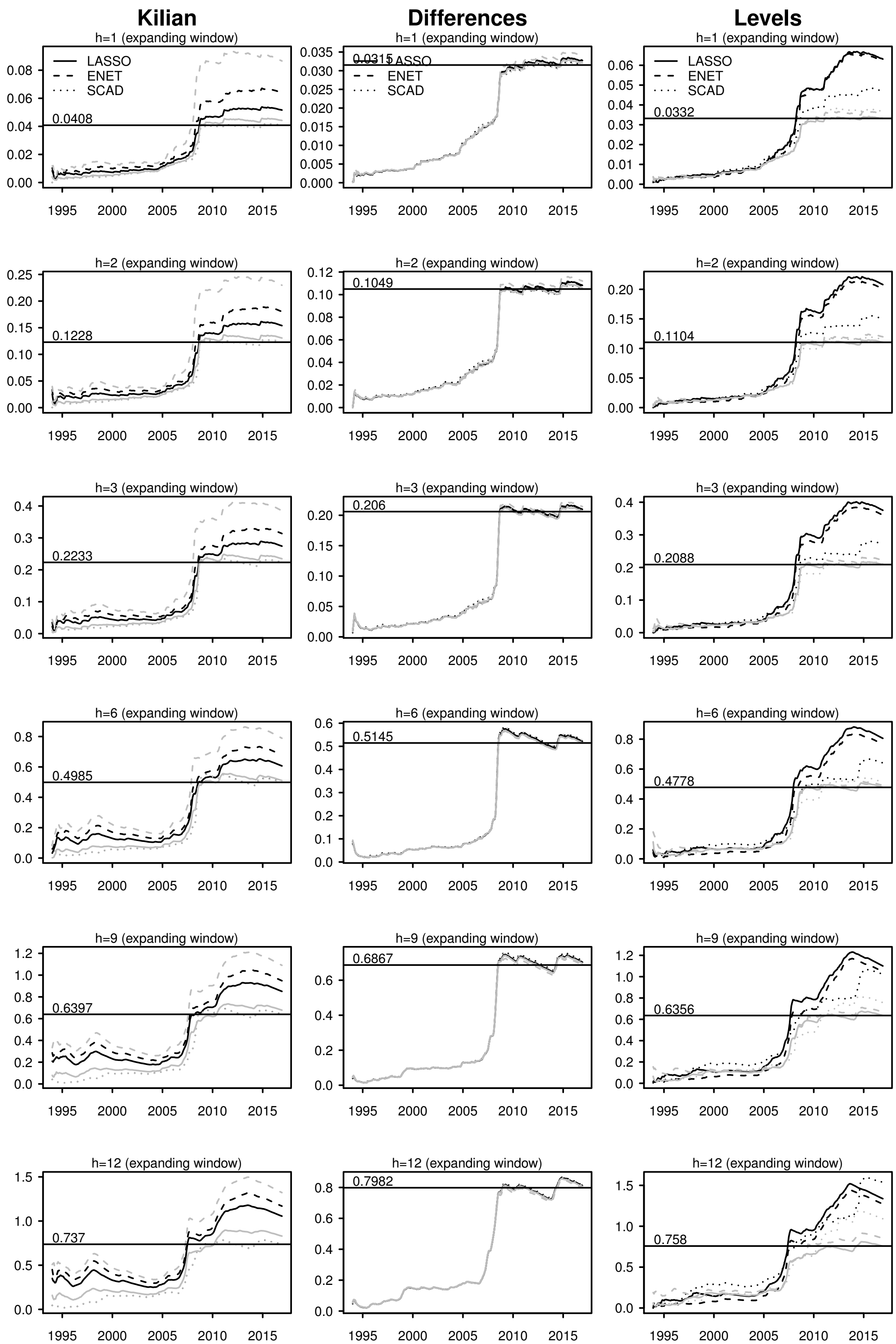

Figure 7.5: Sparse VARs augmented with industrial production (expanding window).

in the previous subsection. In the case of the VAR in differences we observe no improvement and the VAR in differences remains the worst performing model for the longer forecast horizons with larger forecast errors than the Kilian VAR and the VAR in levels. Finally, in the VAR in levels, extended by the log exchange rates, there is a substantial deterioration of forecast performance of all three sparse estimation methods across all forecast horizons. Here now the performance of SCAD also deteriorates.

Taken together, we find no improvement by augmenting the VARs with the production indices or exchange rates and applying the sparse estimation methods to eliminate unimportant variables and lags. There are two aspects leading to this outcome. The first possibility is that the sparse VAR estimators are not able to filter out the relevant variables and lags. Given that, the sparse VAR methods apparently fail to set the parameters to zero which actually are equal to zero leading to more noisy parameter estimates and forecasts, finally resulting in larger MSE values. The second possibility is that the variables used for extending the model are largely irrelevant for forecasting the world oil price or contain information which is already comprised in the three core variables. In the case of the production indices it seems quite plausible that they represent information about economic activity in the G7 countries which is also contained in the global real activity index. This is, however, not born out by the correlations of the production indices with the real activity index (max. correlation ≈ 0.12 with real activity and ≈ 0.38 with changes of real activity). Here it is important to recall that the real activity index is a global measure based on international dry cargo shipping rates and therefore also comprises the activity of other large emerging economies like China and India.

7.4.3 Investment Opportunities

As a third extended variable set we consider the prices of different investment opportunities as possible candidate variables. There is a literature on the stock market effects of oil price shocks (see Kilian and Park, 2009, among many others)(see Kilian and Park (2009) among many others) or on the usefulness of financial market data in forecasting oil prices (e.g. Degiannakis and Filis, 2018). Here, the identifying restrictions imposed postulate the instantaneous response of a stock market index (actually the real log returns of the Center for Research in Security Prices (CRSP) value-weighted market portfolio), but not the other way round. In a VAR there may be, however, also a response of the oil price to the stock market reaction in the next period. Thus, the information comprised in the returns of different investment opportunities may be suitable to improve the oil price forecasts. Those forward-

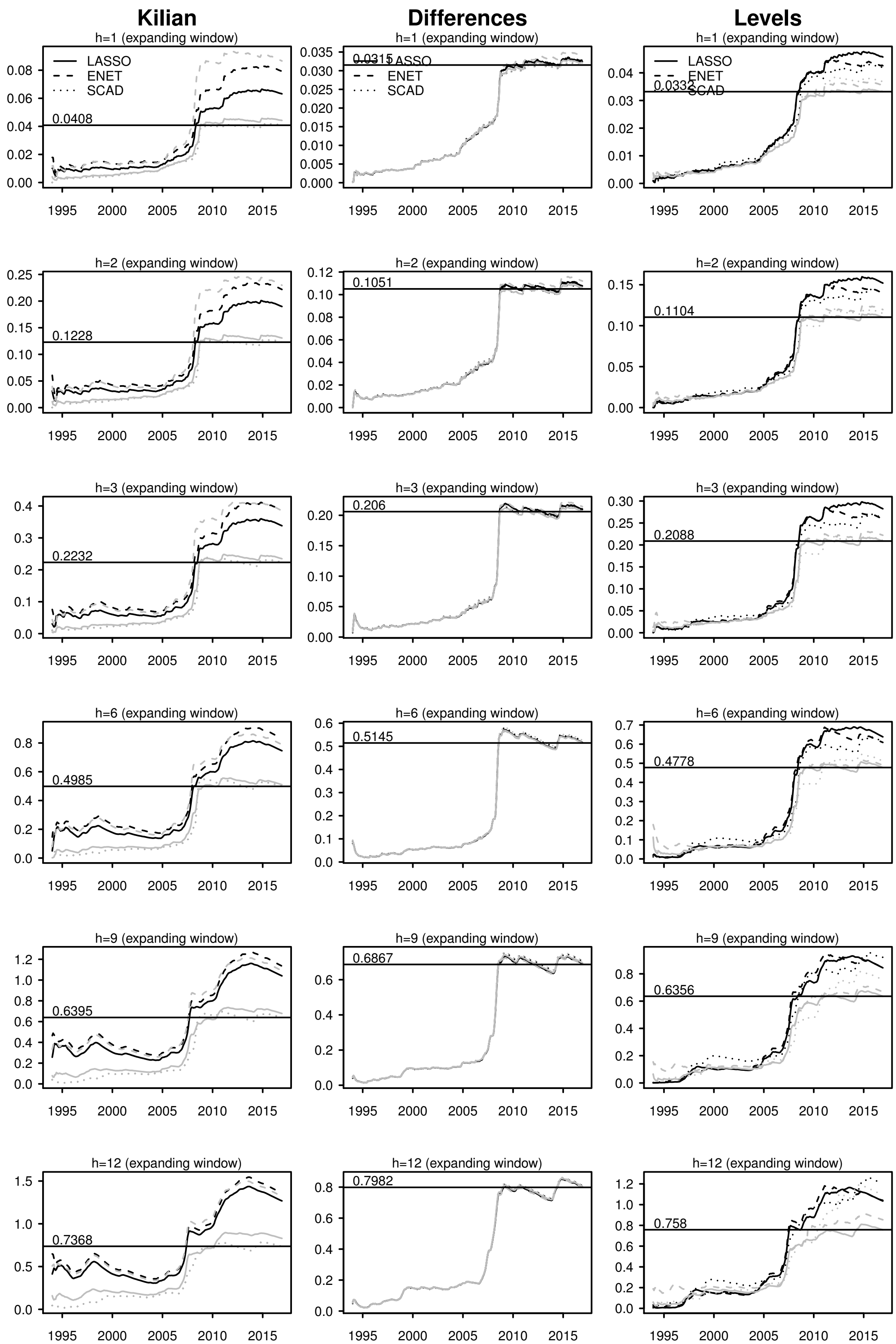

Figure 7.6: Sparse VARs augmented with exchange rates (expanding window).

looking variables are also considered in business cycle research (see e.g. Stock and Watson, 2003).
To assess this issue we include the index values or the returns of CRSP market portfolio[11], the real gold price[12], a comprehensive bond price index[13] in the sparse VAR models. These time series are transformed by logs and are differenced in the cases of the Kilian VAR and the VAR in differences. In addition, the 3 month and 10 year treasury rates[14] are also included without transformation.

The relevance of the additional financial variables is mainly motivated by the arbitrage condition linking the crude oil spot price to crude oil futures prices (see Fattouh et al., 2013). As Hamilton and Wu (2014) (2014) point out, the futures market started to expand very quickly in the early 2000s, primarily due to crude oil futures viewed as an instrument for portfolio diversification. Thus, we include variables that have frequently been used to determine returns in futures markets (see e.g. Bessembinder, 1992; De Roon et al., 2000; Hong and Yogo, 2012).

The results are shown in figure 7.7. As before in the case of the introduction of the production indices and the exchange rates we find no improvement of the forecast performance, in particular since the financial crisis and the great recession. In the Kilian VAR we observe a deterioration of the forecasts based on LASSO and ENET estimates across all forecast horizons. Only SCAD achieves the same performance as the basic model with the three core variables (again shown as gray lines for reference), most likely caused by the total elimination of the additional variables. Considering the VAR in differences we find the forecast errors achieved with the additional variables to be very similar to those without augmentation. The VAR in levels leads to worse forecasts for all augmented models across all horizons. This holds in particular for the SCAD estimated models at larger forecast horizons.

---

[11]This is a broad value-weighted index as the market portfolio, formed on the universe of all CRSP firms incorporated in the US. Data are from the data archive of Kenneth French.

[12]Price of one fine ounce in US$, daily fixed at the London Bullion Market, averaged over the respective month and transformed to a real price by the US CPI. Data are from the time series database of the Deutsche Bundesbank.

[13]This is the BofA Merrill Lynch US Corp Master Total Return Index Value, transformed to a real price by the US CPI.

[14]Both series are retrieved from the FRED database with the corresponding codes GS3M and GS10.

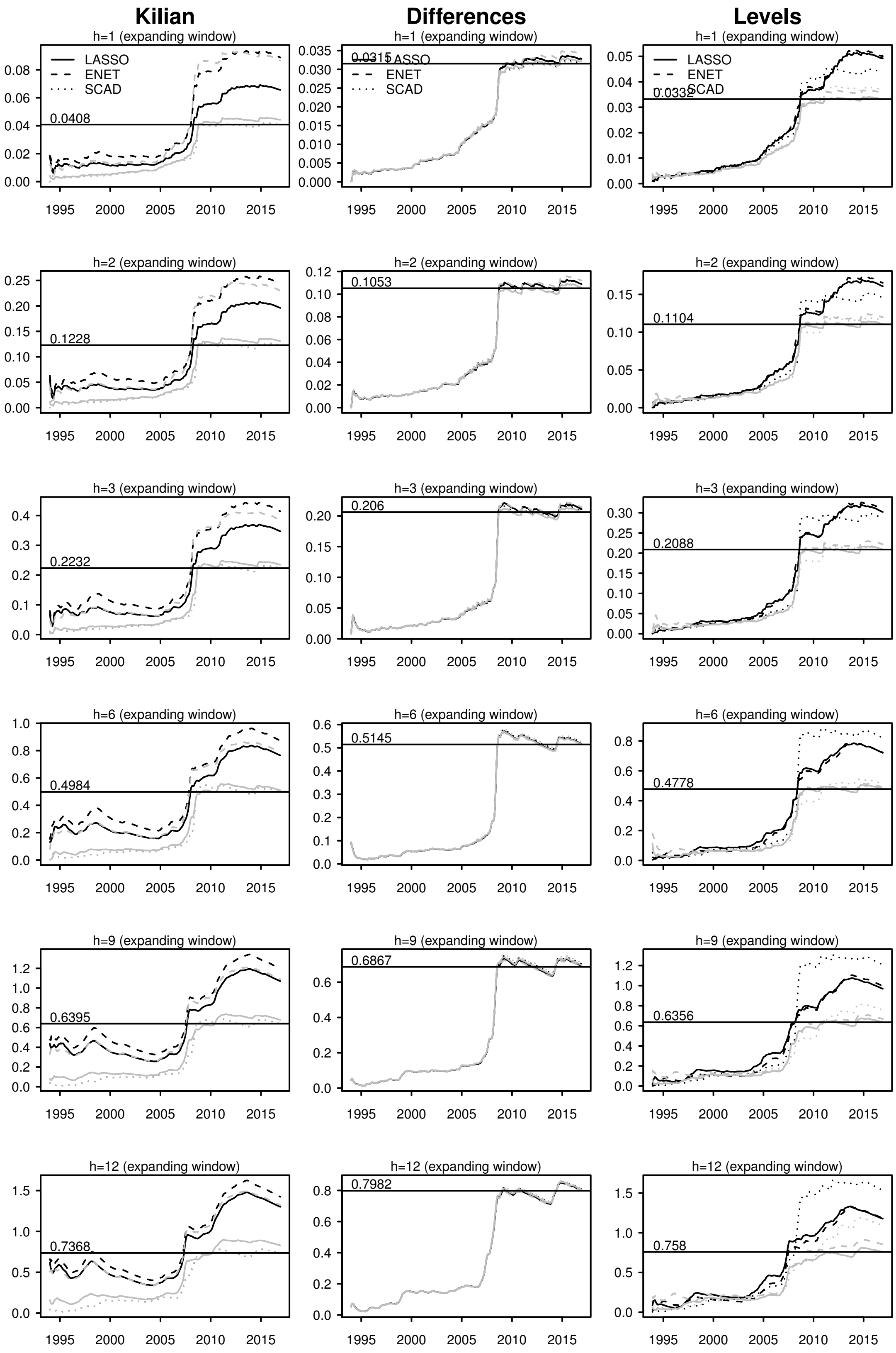

Figure 7.7: Sparse VARs augmented with investment opportunities (expanding window).

7.4.4 Impulse Indicator Saturation

The final attempt to improve the accuracy of the oil price forecasts is the introduction of IIS in the estimation procedure of the VAR. This device, introduced by Santos et al. (2008), uses a complete set of dummy variables (one for each observation) to prune out single observations from the whole data series which may represent outliers or result from structural breaks. Since the complete set of dummy variables cannot be introduced into the VAR at once, the first step of the procedure includes the dummy variables for the first half of the sample period and tests which of them are significant on a 5 percent level. In the second step, only the dummy variables for the second half of the sample period are included and tested in the same way. Finally, all dummy variables which have been found significant in the first and the second step are introduced simultaneously and again individually tested for their significance. The subset of those dummy variables which remain significant in the final step are then kept in the VAR for the estimation. This actually amounts to exclude the associated observations. This procedure is conducted anew in each forecast step of our forecast evaluation procedure.

Figure 7.8 shows the results when all t-tests in the IIS procedure are performed on a 5 percent level of significance. Compared are the VAR with 12 lags estimated by OLS (VAR12) or subjected to the selection of variables by the LASSO, represented by gray lines, with their variants estimated after performing the IIS (denoted IISVAR12 and IISLASSO), represented by black lines. In the case of the VAR(12) the application of IIS does not lead to an improvement of the forecasts and even leads to deteriorations at the longer forecast horizons. This holds likewise for the Kilian transformation and for the VAR in levels, whereas the curves in the case of the VAR in differences are very close.

Comparing the acLASSO-based forecasts with and without the IIS in advance we see that those without IIS nearly always have an edge over those with IIS. Again, the difference becomes larger with increasing forecast horizon for the Kilian transformation and the VAR in levels and is negligible for the VAR in differences. Nevertheless, both LASSO-based forecasts are not far away from the ordinary VAR forecasts before the financial crisis but become much better thereafter. Repeating the analysis with a 1 percent level of significance leads to results (not shown) that are less favorable for the IIS in this application. Finally, turning to a rolling window instead of an expanding window for the estimation and IIS selection leads to considerably larger forecast errors when using IIS (see the appendix).

In sum, we can draw the conclusion of the analysis with the extended information sets in this section that the VAR in levels with just the three core variables estimated by sparse VAR

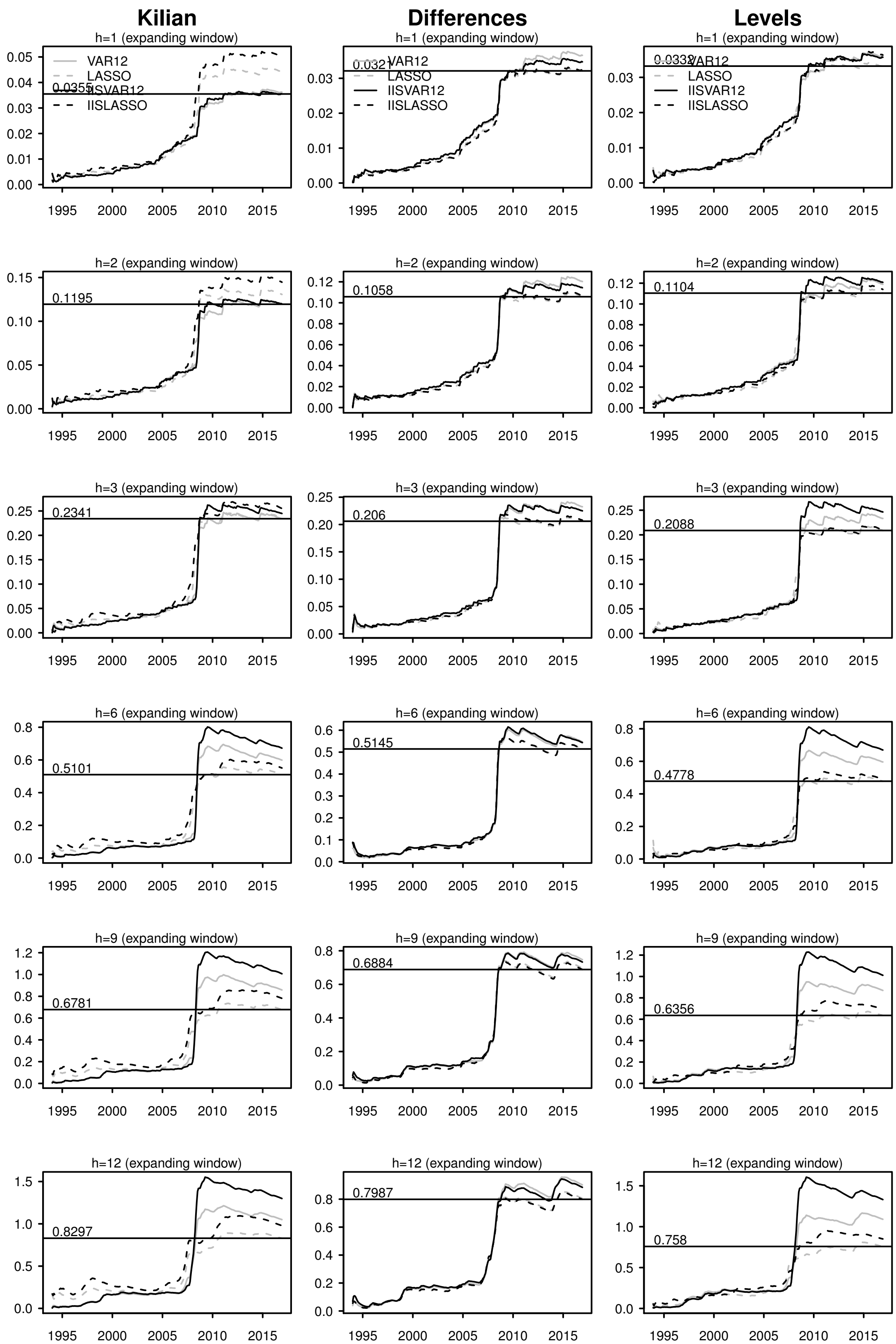

Figure 7.8: Evaluation with IIS (expanding window).

methods (as discussed in section 7.3.2 above) remains the best overall forecast method.

7.5 Conclusion

In the above analysis we have conducted a forecast evaluation exercise for the world real price of crude oil. Our point of departure was the three-variable VAR model of Kilian (2009) which has already been subjected to a forecast evaluation by Alquist et al. (2013). The value added of our analysis is the application of estimation methods based on regularization to achieve sparsity in the parameter matrices of the VARs. Whereas classical information criteria for lag-order selection such as AIC or BIC restrict entire parameter matrices for lag orders higher than the selected to zero, the regularization methods have the property to restrict specific parameters within the parameter matrices to zero while others retain their values different from zero. By that, the detrimental effect of parameter estimates which are truly zero but are estimated with small magnitudes and are likely to be insignificant on forecast performance is reduced. This holds the promise of reaching a better forecast performance.

The main results of this forecast evaluation exercise can be summarized in three main lessons. The first lesson is that the selection of the benchmark VAR reveals that "long" VARs including many lags of the variables (e.g. 12 or 24) has a justification for impulse response analysis, but is detrimental to forecast performance. As emphasized in (Kilian and Lütkepohl, 2017, pp. 63ff.) covering a cycle of a year with monthly data (or a multiple thereof) is important for impulse response analysis. However, we have seen that these "long" VARs are clearly dominated by more parsimonious VARs with respect to forecasting performance. The second lesson is that the application of regularization of the VARs for improving forecast performance depends on the choice of variable transformations. We find that regularization improves forecasts especially for the longer forecast horizons up to 12 months for the VAR with variables transformed according to Kilian (2009) and the VAR in levels. The forecasts for the VAR in differences are also improved by applying the sparse estimators, but here only for the shorter forecast horizons. The third lesson is that extending the variable set and then applying the sparse VAR estimators does not generally lead to further reductions of the forecast errors. Thus, the general property of the LASSO and related estimators as devices for the selection of suitable sets of predictors out of a large set of candidate variables as supposed for simple linear regression is not born out in the case of VARs. Instead, we find that the forecast performance of the augmented VARs worsens or the additional variable-lag combinations are totally pruned out by the sparse estimators.

This outcome stands in contrast to other related macroeconomic forecast evaluation exercises (not focusing on the oil price) such as Nicholson et al. (2017). Therefore, it appears that more experience with the regularization methods for estimating sparse VARs in different situations is required. This paper makes a contribution to this endeavor. Along these lines an investigation of the suitability of other regularization approaches for VARs such as the variants recently proposed by Nicholson et al. (2016, 2017) would be valuable. The current implementation in the R-package “BigVAR” appears rather slow in terms of computation time, however. This may not be a problem for computing a single or a small number of forecasts, but it becomes prohibitive for a forecast evaluation exercise where hundreds or thousands of forecasts need to be computed as we have done in this paper. Hence, we have not applied these methods in the present paper, but look forward to do so in the future when faster computers and/or software are available. Finally, it would be interesting to have a closer look into the estimation results to investigate which variable-lag combinations are eliminated by the regularization and to see whether is pattern is stable over time or is subject to systematic changes. This is beyond scope of the present paper, but is an interesting opportunity for future research.

In the next and final chapter, we will focus specifically on crude oil futures and their forecasting properties given symmetric and especially asymmetric loss.

Chapter 8

Do crude oil futures follow asymmetric market preferences?

Oil price fluctuations have long been associated with U.S. and global business cycles (Hamilton, 1983; Kilian, 2008a, 2009), forecasts of the crude oil prices play, therefore, an important role in forward looking economic models used worldwide by central banks as well as international institutions such as the International Monetary Fund (IMF) to formulate macroeconomic outlooks as well as policy recommendations (Baumeister and Kilian, 2014b). In this regard, futures contracts with maturity h have been commonly used by institutions as the h-period forecast of crude oil prices (Alquist et al., 2013). However, when evaluating the performance of futures as oil price forecasts, empirical evidence strongly suggests their under-performance with regard to econometric models and no-change benchmarks (Alquist et al., 2013).

It is now common knowledge that the spread between commodity futures and spot prices is driven by the so-called risk premium on futures contracts. First proposed by Keynes (1930) in his theory of "normal backwardation", the risk premium on futures is the compensation paid by producers of physical commodities to arbitrageurs in order to hedge their price risk. This should result in futures prices lower than expected spot prices in order to generate a positive expected compensation.

In this analysis we propose a different approach to empirically represent the risk premium. We treat crude oil futures as forecasts of the spot price of crude oil. The crude oil futures market acts as forecaster with an underlying loss function that is potentially asymmetric. If the theory of normal backwardation applies to crude oil futures, a positive risk premium

This chapter is based on joint work with Julian LeCrone.

should be reflected by a loss function that indicates an aversion to falsely overestimating the spot price of oil (i.e. future prices exceeding the actual spot price of oil in h periods). To analyze if futures are forecasts following a symmetric loss function, we first apply the (modified) Mincer and Zarnowitz (1969) framework to test for unbiasedness and efficiency. Then, to quantify the possible degree of asymmetry, we use the approach proposed by Elliott et al. (2005, 2008). To control for structural breaks in the data, we follow Hamilton and Wu (2014) by splitting the sample in 2005 when they identify a structural break in the evolution of the risk premium on crude oil futures. Furthermore, we apply step-indicator saturation (SIS) as developed by Castle et al. (2015) to detect and estimate structural shifts in the spot price of crude oil as a second approach to verify the robustness of our results.

The remainder of the chapter is organized as follows: Section 1 introduces the concept of loss functions in the forecast evaluation literature and presents a review of the literature on symmetric and asymmetric loss, on the risk premium in oil futures markets and why failing to account for it results in biased estimates of market expectations. In section 2 we develop the estimation framework that is applied to the crude oil spot prices and futures. Section 3 presents and discusses the results. The concluding remarks follow in section 4.

8.1 Motivation, theory and formal notation

In this section, at first the circumstances under which oil futures prices can be seen as market expectations, or more precisely, as market forecasts of the spot price of oil are explained. Then, the theoretical literature about the so called risk premium on futures is presented and discussed. The understanding of the connection between crude oil prices and futures prices with maturities between $h = 1$ and $h = 6$ months as well as the possibility of a risk premium is fundamental in this context. Finally, the notation used throughout the chapter is presented and the concept of a loss function is introduced.

8.1.1 The risk premium in oil futures markets

Oil futures contracts allow a producer of the commodity to fix today the price and stipulate the delivery of the commodity in question in h months. The contract may then be re-traded on a futures market such as the New York Mercantile Exchange (NYMEX) between inception and maturity. Futures trading thus allows for better risk allocation and, more importantly, for new information about the expected spot price of oil exchanged through futures price adjustments as trade of contracts continues (Grossman, 1977; Danthine, 1978). Baumeister and Kilian (2016b) point out that oil futures with maturity h are widely used by policymakers as a measure of market expectations (forecasts) regarding the crude price of oil in h periods. However, oil futures are characterized by a premium on top of the expected price of crude oil. In simple terms, the risk premium is the expected positive return that investors expect from an investment in oil futures contracts (Fama and French, 1987; Grossman, 1977). This leads to the conclusion that using futures as forecasts for the spot price of oil is only valid if the risk premium is negligible.

Let y_{t+h} be the realization of the spot price in period $t+h$ and f_{t+h} the price of an h-months futures contract in period t. In order to extract the correct market signal regarding the spot price of oil, the unobservable risk premium RP_t paid to the long side has to be accounted for:

$$y_{t+h} = f_{t+h} + RP_t + \epsilon_t. \tag{8.1}$$

The stochastic term ϵ_t comprises unexpected price disturbances with $E(\epsilon_t) = 0$. Transforming equation (8.1) allows to estimate the risk premium, using the following linear regression model:

$$y_{t+h} - f_{t+h} = RP_t + \epsilon_t = \boldsymbol{\beta}_h' \boldsymbol{x}_t + \epsilon_t \tag{8.2}$$

$\boldsymbol{\beta_h}$ being the regression coefficients for futures contract maturity h in this equation, the vector $\boldsymbol{x}_t$ contains the predictors and ϵ_t is again the zero mean error term. Empirical evidence for the existence of a positive risk premium in commodity markets, crude oil markets in particular, was found in early studies such as Chang (1985) or Fama and French (1987). While the choice of the explanatory variables relies on the literature developed on risk premium in commodity and oil futures markets, there is no consensus which variables to include in $\boldsymbol{x}_t$.

Among others, regularly employed explanatory variables are macroeconomic risk variables, such as unexpected CPI inflation, US T-bill rates or the term structure of interest rates (e.g. Bessembinder, 1992; Sadorsky, 2002; Pagano and Pisani, 2009). Similarly, equity and bond market variables such as the return on stock market indices or yield spreads (e.g. Bessembinder and Chan, 1992; De Roon et al., 2000; Hong and Yogo, 2012) are used. Moreover, some authors suggest for crude oil futures the use of industry specific explanatory variables such as changes in crude oil inventories or expected default frequencies in the oil and gas industry (e.g. Gorton et al., 2012; Acharya et al., 2013). The diversity of proposed variables has resulted in a wide range of risk premium estimates for the oil futures market.

More recently, Hamilton and Wu (2014) use a term structure model to estimate the risk premium on crude oil futures and to find evidence for a time-dependent risk premium. Whereas before 2005 the long-term position of future contracts received stable and positive returns for assuming the price risk from sellers, they observe a much more volatile and decreasing evolution of the risk premium since 2005. They attribute such shift from the long to the short side of futures contracts to increasing participation of "index-funds" as buyers in oil futures markets.

As Singleton (2014) points out, the empirical evidence for a risk premium in oil futures markets is compelling. Baumeister and Kilian (2016a) evaluate risk-premium-corrected oil futures as forecasts for the spot price of crude oil, estimating the risk premium $\widehat{RP}_t$ resulting from a wide range of models as described above, and construct market expectations as follows: $E\left(y_{t+h}\right) = f_{t+h} + \widehat{RP}_t$. After ranking the risk-premium-corrected futures as forecast for the spot price of oil by computing their mean squared percentage error (MSPE), they conclude that the Hamilton and Wu (2014) risk premium adjusted oil futures perform best in comparison with other alternatives. Evidence in the literature allows the conclusion that the use of futures as market forecasts of the spot price of oil is invalidated by the existence of a non-zero risk premium. This risk premium is however poorly understood and quantified.

A closer look at crude oil spot prices as well as predictions from oil futures in figure 8.1 provides some useful insights: In the upper plot, the black curve shows the average monthly spot price of WTI crude oil over the period 1983-2017. For each point in time t the red dots represent the h-months crude oil futures prices for $h = 1, ..., 6$.

Visually, the sample period can be split into two subsamples. A first subsample that roughly corresponds to the period between 1983 and 2002, in which price movements on spot markets remained limited. The oil price collapse in 1986 as well as the temporary price spike between 1990 and 1991 are two notable exceptions. Visually, the futures prices can roughly be interpreted as no-change forecasts during this first period. Starting in 2002, an increasing price trend emerges on crude oil markets. However, prices of the corresponding futures seem to indicate that the market did not anticipate this trend. After a price crash, following the collapse of Lehman Brothers, and the subsequent financial and economic crisis in 2008, the price remained stable for around four years, before decreasing from around \$100 in 2014 to \$30 beginning of 2016. Again, the futures market failed to recognize the developments of the

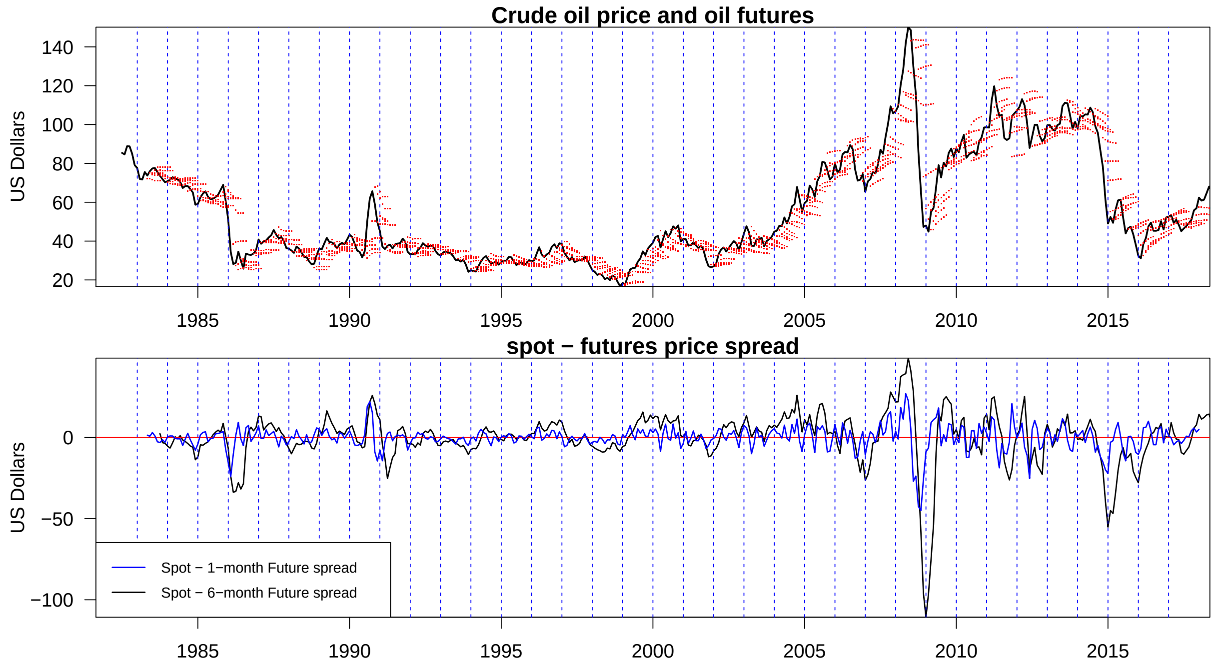

Figure 8.1: Crude oil spot price (WTI benchmark), futures prices as traded on the NYMEX and the implied spot - futures price spread 1983-2017 in constant 2017 \$ (Source: FRED, Thomson Reuters).

spot prices, even when looking at 1-month futures prices.

The lower plot of figure 8.1 shows the spot price vs. futures price spread for maturities $h = 1$ and $h = 6$. This spread is also interpreted as the forecast error implied by the different futures maturities i.e. $e_{t+h} = y_{t+h} - f_{t+h}$. The blue curve corresponds to a 1-month futures spread e_{t+1} and the black line to a 6-months futures spread e_{t+6}. In both cases similar observations can be made: During the first subsample period from 1986 to 2002, the variability in the spread remains limited for both the 1-month and the 6-months futures, with the exception of the two before-mentioned events, which resulted in stronger and longer-lasting spreads. More important is, however, the evolution in the second subsample: From 2002 onwards the volatility in the spread increases significantly and is more pronounced for higher futures maturities. Figure 8.1 illustrates that relying on oil futures prices as forecasts for the price of crude oil, as is commonly done by central banks and international organizations (Baumeister and Kilian, 2014b), can be highly misleading as an indicator of market expectations of the spot price of oil.

In the following, irrespective of such limitations, crude oil futures are treated and used as market forecast of the spot price of oil, under the assumption of an underlying asymmetric loss function. Consequently, a brief introduction of a few concepts concerning the formal notations used throughout the chapter as well as the concept of symmetric and asymmetric loss are given below.

8.1.2 The concept of symmetric and asymmetric loss function

As already stated, y_{t+h} stands for the realization of the target variable of interest, i.e. the spot price of oil in period $t+h$, f_{t+h} denotes the forecast of this variable in period t, i.e. the price of a crude oil futures with maturity h in period t. The market is assumed to act as forecaster of the crude oil price through futures trading, prices being the result of supply and demand: Sellers and buyers agree on the equilibrium price based on their available information. Formally, market participants use information available to them in period t and this information is summarized in the information set Ω_t. The difference between realization and forecast of the target variable is the forecast error $e_{t+h} = y_{t+h} - f_{t+h}$. It is important to distinguish between a positive forecast error ($e_{t+h} > 0$) and negative forecast errors ($e_{t+h} < 0$). The first case results in an under-prediction ($f_{t+h} < y_{t+h}$), in the latter case an over-prediction ($f_{t+h} > y_{t+h}$) of the target variable of interest.

The way that agents, which produce forecasts, use and incorporate the information available to them is a frequent and relevant topic in economic research. A fundamental question is whether forecasters are rational in their information processing. If rationality is rejected by statistical testing, there exists the possibility to improve forecasts. For this purpose the introduction of the concept of the loss function is required. First used by Granger (1969), it describes the economic cost or loss induced by a forecast error. Given the information set available at time t, a loss function $L(.)$ maps the realizations and the forecasts to non-negative real numbers so that $L(y_{t+h}, f_{t+h}|\Omega_t) \in \mathbb{R}_+$. Alternatively, the loss function can also be specified as a function of the forecast error, so that $L(e_{t+h}|\Omega_t) \in \mathbb{R}_+$.

As Elliott et al. (2008) note, most empirical work in testing forecast rationality has focused on the assumption of MSE loss, defined as $\left(MSE = \frac{1}{T}\sum_{i=1}^{T} e_{t+h}^2\right)$. Another loss function that is frequently used because it has the same scale as the data is the relative mean squared error (RMSE) $\left(RMSE = \sqrt{\frac{1}{T}\sum_{i=1}^{T} e_{t+h}^2}\right)$ (Hyndman and Koehler, 2006). Both, MSE and RMSE have in common that they belong to the family of symmetric loss functions, making the forecaster indifferent between an under- or an over-estimation of the target variable: A positive forecast error e_{t+h} has the same loss as a negative forecast error of the same magnitude $-e_{t+h}$ so that $L(e_{t+h}|\Omega_t) = L(-e_{t+h}|\Omega_t)$.

The assumption of symmetrical loss was questioned by Granger (1969) who argued that the cost implied by a positive or negative forecast error should be different. Even most primitive economic conditions should result in an asymmetric loss function (Elliott et al., 2008). This implies that the loss induced by a positive forecast error is different from the one associated with a negative forecast error of them same magnitude $L(e_{t+h}|\Omega_t) \neq L(-e_{t+h}|\Omega_t)$. This simple difference confirms that it might be rational for a forecaster to systematically bias his forecasts. If e.g. the cost induced by a positive forecast error (under-prediction) is higher than the one of a negative forecast error (over-prediction), a forecaster facing uncertainty will have an incentive to reduce the occurrence of under-predictions in comparison with over-predictions, thus creating a bias. Biased forecasts might, therefore, be a first indication for asymmetric loss (Keane and Runkle, 1990).

In the case of crude oil futures, one can assume that the oil futures market, acting as forecaster of the spot price of oil, makes its forecasts on the basis of an unobservable loss function, resulting from supply and demand dynamics as well as underlying preferences of market participants. In making its forecasts, the market implicitly minimizes the loss. The most intuitive way to consider the possibility of asymmetric loss in the oil futures market is to adopt the point of view of a potential investor. Whereas the information available to an investor on the future state of the oil market might be imperfect, he receives a strong incentive to minimize informational deficits in order to form the best possible expectations regarding future spot prices.

If we suppose that the buyer of oil futures is interested in a positive return on investment, he will only buy those futures contracts that have a price below his expected price of crude oil in h periods, so that the difference (an expected positive forecast error) compensates him for the price risk he has to bear. If futures prices are higher than the expected price of oil in h-months, no trade will occur, because buyers would expect a loss on their investment (negative forecast error). Thus, futures prices fall until they reach a level below the expected price of crude oil (from the investor's perspective) in correspondence with the required risk premium.

It is worth noting that in the short term, the market entry of index funds, increasingly the case since 2005 (Hamilton and Wu, 2014), diminishes the risk premium and, thus, the expected asymmetry. This seems due to the fact that index funds always wish to hold a portfolio containing a certain mix of futures contracts with different maturities. As Hamilton and Wu (2014) note, the participation of index funds can even change the recipient of the risk premium from the long to the short side of the contract, meaning that from a buyers perspective the risk premium becomes negative. The reduction of the risk premium would also be the case, when a buyer of oil futures is not interested in positive returns (such as classical arbitrageurs), but in hedging against rising crude oil prices. The risk premium should then depend on the risk aversion of market participants with respect to the expected value of the spot price.

8.2 Methodological framework

In the framework of this study, an alternative view on the risk premium in oil futures markets is proposed. The assumption of incorrect predictions by futures evaluated on the basis of a symmetric loss function, such as the MSE or RMSE would be erroneous, if the underlying loss proves to be asymmetric. The existence of a non-negative risk premium, implied by the literature, is a first indication of asymmetric loss.

An introduction is presented on how to test for biasedness and efficiency under symmetric loss, in order to identify indicators for asymmetry in the loss function. Structural breaks are considered in testing for unbiasedness and efficiency. step-indicator saturation (SIS) is presented as a method used to estimate structural breaks in time series, which are consequently controlled within the Mincer-Zarnowitz (MZ) framework. In the second part of this section we introduce the flexible loss function developed by Elliott et al. (2005, 2008) that allows for asymmetric loss.

8.2.1 Unbiasedness and efficiency under symmetric loss

The Mincer and Zarnowitz (1969) regression is typically used to test for forecast unbiasedness under symmetric and quadratic loss:

$$y_{t+h} = \beta_0 + \beta_1 f_{t+h} + \epsilon_t \tag{8.3}$$

Where the joint null hypothesis for unbiasedness is H_0: $\beta_0 = 0, \beta_1 = 1$.

If the null hypothesis can be rejected, there is evidence for biased forecasts, that would translate into a systematic over- or under-prediction implied by oil futures. Because of non-stationary regarding the time series of the spot and futures oil prices, equation (8.3) can be modified in the following way by enforcing the restriction $\beta_1 = 1$ and subtracting the forecast f_{t+h} on both sides of the equation:

$$e_{t+h} = \beta_0 + \epsilon_t \tag{8.4}$$

Where e_{t+h} is the forecast error observed in period $t + h$, implied by the h-month futures contract traded in period t. In case of equation (8.4) the null hypothesis H_0: $\beta_0 = 0$ tests, whether the forecasts are systematically biased.

As an extension of the MZ approach to evaluate for unbiasedness, forecast efficiency can be tested by augmenting equations (8.3) and (8.4) with $\boldsymbol{w}_t$, a subset of variables, selected from the information set Ω_t available to market participants. This results in following two equations:

$$y_{t+h} = \beta_0 + \beta_1 f_{t+h} + \boldsymbol{\beta}'_{2h}\boldsymbol{w}_t + \epsilon_t \tag{8.5}$$

$$e_{t+h} = \beta_0 + \boldsymbol{\beta}'_{1h}\boldsymbol{w}_t + \epsilon_t \tag{8.6}$$

To evaluate whether market participants use the information available efficiently, one can test the following two null hypotheses H_0: $\boldsymbol{\beta}_{2h} = 0$ and H_0: $\boldsymbol{\beta}_{1h} = 0$ respectively. If the null hypothesis is rejected, there is evidence in both cases that variables in $\boldsymbol{w}_t$ contain useful information that can be used to reduce the forecast error, when a symmetric loss function is correct. As with biasedness, inefficiency in the symmetric case can also be an indication for an asymmetric loss function.

8.2.2 The Mincer-Zarnowitz framework in the presence of structural breaks

Recent research indicates that employing the MZ-framework without taking into account the presence of structural breaks might lead to a false rejection of unbiasedness. Using a MZ framework based on monthly data, Capistrán (2008) analyses the inflation forecasts of the FED and concludes that they are biased. He contradicts Romer and Romer (2000) and Sims (2002) that previously came to the conclusion of unbiased forecasts by the FED. While Romer and Romer (2000) and Sims (2002) do not control for structural breaks in their

MZ-framework, Capistrán (2008) identifies and incorporates structural breaks into the estimation. The appointment of Paul Volcker as chairmen in 1978 changed the FED's forecast behavior from a systematic under-prediction (positive errors) in the sub-sample until 1987 into a systematic over-prediction (negative errors) in the second sub-sample. Thus, both effects cancel out in the Romer and Romer (2000) study and resulting in falsely rejecting unbiasedness.

Another method to control for structural changes while evaluating forecast rationality would be to split the data in subsamples (e.g. Croushore, 2012; Patton and Timmermann, 2012) and apply common testing methods to the subsamples. However, as Rossi and Sekhposyan (2016) note, in many cases the choice of subsamples might not be as straight forward as some better understood economic events, such as the Great Moderation or the Great Recession. Furthermore, when multiple breaks exist in the data or the break dates are unknown, additional robustness checks, based on different subsamples, might be necessary. Thus, Rossi and Sekhposyan (2016) propose an unbiasedness and efficiency testing framework that is robust with respect to structural instabilities. In an empirical application to the FED's Greenbook forecast they confirm the main results of Capistrán (2008) and Patton and Timmermann (2012) and find significant evidence against the FED's forecast rationality as previously stipulated by Romer and Romer (2000) and Sims (2002).

Using a similar approach Sinclair et al. (2010) test whether the FED incorporates information about the business and the inflationary cycle into its forecasts of GDP growth and inflation respectively. For current period forecasts the information is taken into account, however, this is not the case for one-quarter ahead forecasts. When the state of the economy is not taken into account, the FED produces inefficient forecasts. The inefficiency disappears once the state of the economy is known to the FED.

Because of the fact, that the oil market has historically been characterized by industry-specific structural and technological changes (Hamilton, 2011), as well global shifts in demand due to global business cycle patterns (e.g. Hamilton, 1983; Kilian, 2009; Kilian and Murphy, 2014) it is necessary to account for these shifts when testing for unbiasednesss and efficiency in the MZ-framework. Therefore, we propose following extended MZ-equation to test for unbiasedness and efficiency similar to equation (8.6):

$$e_{t+h} = \beta_0 + \boldsymbol{\beta}'_{1h}\boldsymbol{w}_t + \boldsymbol{\beta}'_{2h}\boldsymbol{s}_t + \epsilon_t. \tag{8.7}$$

For testing unbiasedness and efficiency the null hypothesis remains the same as in the case already mentioned: $\mathbf{H}_0$: $\beta_0 = 0$ and $\mathbf{H}_0$: $\boldsymbol{\beta}_{1h} = 0$. This approach to test for forecast unbiasedness and efficiency in futures forecasts controls for structural breaks via two different specifications. First, the structural break in 2005 identified by Hamilton and Wu (2014) is included, secondly, a set of structural breaks in the crude oil spot market, as estimated by SIS, is taken into consideration, as described below.

8.2.3 Structural break identification based on Step-Indicator Saturation

Ignoring structural breaks in tests for forecasting rationality with the classical MZ equations, might lead to false evidence concerning unbiasedness and efficiency, when over the full sample systematic over- and under-estimations average out (see Capistrán, 2008; Rossi and Sekhposyan, 2016). The identification of structural breaks in the crude oil spot and futures markets relies on the work done by Santos et al. (2008), Hendry and Doornik (2014) and Castle et al. (2015) on indicator saturation methods in the context of model selection.

Closely related to impulse-indicator saturation (IIS) presented and used in the previous chapter, Castle et al. (2015) propose and analyze a more powerful approach applied to time series in order to identify structural shifts in the expected level of a dependent variable. Instead of single impulses for each time period, they add a complete set of step indicators to a linear regression model:

$$e_t = \beta_0 + \boldsymbol{\beta}_1' \boldsymbol{w}_t + \sum_{j=1}^{T-1} \delta_j 1_{\{t \leq j\}} + u_t \quad for \quad t = 1, ..., T \tag{8.8}$$

y_t being the dependent variable, $\boldsymbol{w}_t$ encompassing all relevant explanatory variables and the indicator function $1_{\{t \leq j\}} = 1$ for observations up to period j and zero thereafter. Because of the fact that the number of variables to be estimated exceeds the number of observations, it is necessary to split the regression approach in the same way as with IIS.

In a first regression step, Castle et al. (2015) propose to add the first $T/2$ indicators to the linear model and retain those with significant coefficients at the α significance level. In the next regression step, the first $T/2$ indicators are dropped and the second block of indicators is added to the model, again retaining the significant ones. Finally, in a third regression all retained variables (if any) are included and the significant ones are identified as structural shifts in the expected level of the dependent variable. The detection of location shifts in constant conditional models by SIS has furthermore the correct null retention frequency for a nominal selection size of α.

Empirically IIS and SIS have been employed in diverse research fields, among others by Stillwagon (2016) to identify and evaluate structural shifts in exchange rate markets, by Pretis et al. (2015) to evaluate shifts in climate change models. Castle et al. (2016) examine forecast failure in the presence of structural breaks, as illustrated by UK GDP growth and unemployment rate forecasts.

8.2.4 GMM estimation

In order to determine whether a forecaster has asymmetric preferences Elliott et al. (2005, 2008) introduce a flexible loss function:

$$L(e_{t+h}; \alpha, p) = [\alpha + (1 - 2\alpha)\mathbf{1}(e_{t+h} < 0)] \cdot \left|e_{t+h}\right|^p. \tag{8.9}$$

With $0 < \alpha < 1$ being the parameter representing the degree of asymmetry (not to confound with the significance level) and $p > 0$ the one determining the loss function curvature. Furthermore $\mathbf{1}(.)$ is the indicator function that takes the value of one, whenever the condition $e_{t+h} < 0$ holds true and is equal to zero else. The indicator function in conjunction with α assigns different costs to over- or under-predictions of the same value in the loss function. When fixing the value of the asymmetry parameter to $\alpha = 0.5$ one recognizes the special case of a symmetric loss function. In the case of $p = 1$ the loss corresponds to the mean absolute error (MAE) and in the case of $p = 2$ to the MSE. Allowing for $\alpha \neq 0.5$ results in a piecewise linear (Lin-Lin) and in a piecewise quadratic (Quad-Quad) loss function. The interpretation of α is as follows: a value $\alpha < 0.5$ implies that under-predicting the target variable ($e_{t+h} > 0$) is associated with a lower loss than an over-prediction ($e_{t+h} < 0$). Figure 8.2 shows Lin-Lin (p=1) and Quad-Quad (p=2) loss functions with an asymmetry parameter $\alpha = 0.3$ and their symmetric counterparts MAE and MSE.

In both cases the parameter $\alpha < 0.5$ results in a clock-wise rotation of the loss curves. It becomes evident that in comparison with the symmetric cases, the asymmetric loss functions assign different costs to positive and negative forecast errors of the same magnitude. In the case of $\alpha = 0.7$ the cost that results from a positive forecast error ($e_{t+h} > 0$) is higher than the cost resulting from a negative forecast error of the same magnitude. A forecaster with such a loss function will have the tendency to over-estimate the target variable ($y_{t+h} - f_{t+h} < 0$).

It is important to note that even small deviations from symmetric loss ($\alpha = 0.5$) imply

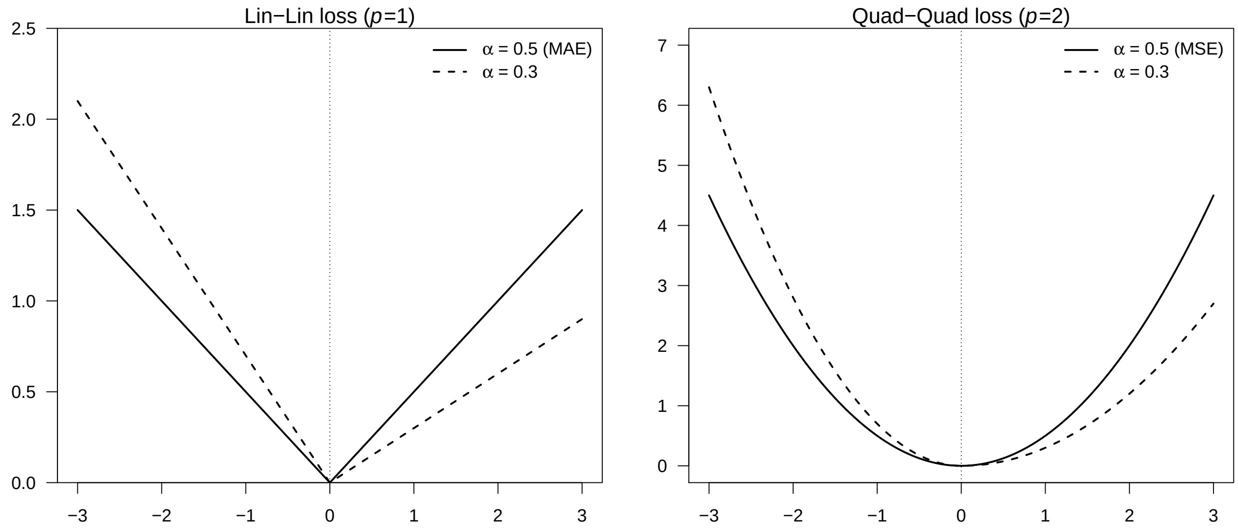

Figure 8.2: Different symmetric and asymmetric loss functions.

rather important loss differences. Take the example of $\alpha = 0.45$: It implies a loss ratio of positive and negative forecast errors of $\alpha/(1-\alpha) = 0.45/0.55 \approx 0.82$ corresponding to an approximate loss difference of 18 per cent. This underscores that even such an important deviation from symmetry would be hardly detected by testing the null hypothesis $\alpha = 0.5$ on commonly used significance levels. Thus, in the following, estimation results regarding α have to be interpreted based on their statistical as well as their economic significance. A small, insignificant deviation from the symmetric case might already be an indication for considerable asymmetric preferences of the forecaster.

The fundamental idea behind the EKT approach is to estimate α while fixing p (or to estimate both parameters simultaneously) according to the following first-order condition:

$$E\left(L'(e_{t+h};\alpha,p)|\Omega_t)\right) = 0 \tag{8.10}$$

with $L'(.)$ being the derivative of the loss function with respect to e_{t+h}. When condition (8.10) is satisfied optimality is achieved, i.e. there remains no information in the information set Ω_t to further decrease the forecast errors. Optimality can then be tested using the following orthogonality condition:

$$E\left(\boldsymbol{w}_t \cdot [\alpha + \mathbf{1}(e_{t+h} < 0)] \cdot \left|e_{t+h}\right|^{p-1}\right) = 0 \tag{8.11}$$

where $\boldsymbol{w}_t$ is a $d \times 1$-dimensional subset of instrumental variables belonging to the information set Ωt and known in period t, when the forecast is made. The optimality (rationality) test is carried out by applying GMM estimation of the parameter α (while holding p fixed) and computing the usual J-statistic for model over-identification to test the null hypothesis of validity (Hansen, 1982).

Furthermore, the EKT approach allows to assess whether the forecaster used the information available in a rational way which becomes evident when deriving the first-order conditions needed for the GMM estimation. If no more information remains in the set of instrumental variables, one should be able to reject the following null hypothesis $\mathbf{H}_0 : \beta = 0$ in $L'(e_{t+h}; \alpha, p) = \boldsymbol{\beta}'\boldsymbol{w}_t + u_{t+h}$. Consequently the moment conditions $E\left(\boldsymbol{w}_t \cdot u_{t+h}\right) = 0$ can be rewritten as $E\left(\boldsymbol{w}_t \cdot \left(L'(e_{t+h}; \alpha, p) - \boldsymbol{\beta}'\boldsymbol{w}_t\right)\right) = 0$ which then reduces to equation (8.11) under the null of rationality.

The EKT approach reverses the methodology in comparison with a traditional and modified Mincer and Zarnowitz (1969) framework in order to determine whether a forecaster is rational and produces optimal forecasts. In their approach, the estimation of the shape parameter α, while fixing $p = 1$ and $p = 2$, leads to a loss function that is most compatible with the assumption of rational forecasters, using all available information efficiently. On the other hand, the traditional approach proposes an assumed loss function, usually the symmetric MSE loss, and then uses the MZ equation to test for unbiasedness and efficiency of forecasters.

To test the assumption of information efficiency that has to coincide with the optimal value of zero in the GMM estimation of the target function, Elliott et al. (2005) propose the following J-statistic for over-identifying restrictions:

$$
\begin{aligned}
J = T & \left\{ \frac{1}{T} \sum_{t=0}^{T-1} \boldsymbol{w}_t \left[\hat{\alpha} - \mathbf{1}\left(e_{t+h} < 0\right) \right] \cdot \left| e_{t+h} \right|^{p-1} \right\}' \times \hat{\boldsymbol{S}}^{-\mathbf{1}} \\
& \times \left\{ \frac{1}{T} \sum_{t=0}^{T-1} \boldsymbol{w}_t \left[\hat{\alpha} - \mathbf{1}\left(e_{t+h} < 0\right) \right] \cdot \left| e_{t+h} \right|^{p-1} \right\}
\end{aligned}
\tag{8.12}
$$

with $\hat{\boldsymbol{S}}$ a consistent estimator of $\boldsymbol{S} = E\left(\boldsymbol{w}_t \boldsymbol{w}_t' \left[\mathbf{1}\left(e_{t+h} < 0\right) - \alpha \right]^2 \left| e_{t+h} \right|^{2p-2} \right)$ and $\hat{\alpha}$ being a linear instrumental variable (IV) estimator of α as follows:

$$
\hat{\alpha} = \frac{\left[\sum_{t=0}^{T-1} \boldsymbol{w}_t \left| e_{t+h} \right|^{p-1} \right]' \hat{\boldsymbol{S}}^{-\mathbf{1}} \left[\sum_{t=0}^{T-1} \boldsymbol{w}_t \left| e_{t+h} \right|^{p-1} \mathbf{1}\left(e_{t+h} < 0\right) \right]}{\left[\sum_{t=0}^{T-1} \boldsymbol{w}_t \left| e_{t+h} \right|^{p-1} \right]' \hat{\boldsymbol{S}}^{-\mathbf{1}} \left[\sum_{t=0}^{T-1} \boldsymbol{w}_t \left| e_{t+h} \right|^{p-1} \right]}
\tag{8.13}
$$

Given $d > 1$ instruments used in the estimation, and due to the fact that one degree of freedom is lost in the estimation of α, the J-statistic is asymptotically χ^2_{d-1} distributed under the null of optimality. As mentioned the whole procedure consists of a joint estimation of the shape parameters of the loss function as well as a test of forecast optimality. Specifically, given the instrumental variables $\boldsymbol{w}_t$ the EKT approach estimates the value $\hat{\alpha}$ while fixing $p = 1$ and $p = 2$, that are consistent with forecast optimality. Consequently, if no optimal loss parameter values exist because some moment conditions result in considerable different estimates, the J-statistic will reject the null hypothesis of optimality. For the purpose of the empirical analysis that follows, if optimality is rejected by the J-test for over-identification we do not consider the value of $\hat{\alpha}$ as valid.

8.3 Empirical application

This section provides an application of the before-mentioned methods to crude oil spot and futures prices. First, the instrument sets, data used and their sources are presented. Then SIS is applied to the WTI crude oil price series, in order to detect structural level shifts that need to be accounted for. Based the results achieved, a review concerning unbiasedness and efficiency in presented, using the MZ framework under symmetric loss, before using the EKT approach to estimate and evaluate the degree of asymmetry in the crude oil futures market. Finally, an exemplary rolling window estimation of the asymmetry parameter $\hat{\alpha}$ is presented in order to evaluate time-varying market preferences.

For all specifications, the unbiasedness and efficiency hypotheses are tested by standard F-tests based on a heteroscedasticity and autocorrelation consistent (HAC) covariance matrix estimator (Newey and West, 1987). Relaxing the condition of symmetry, the GMM estimation follows the continuously updating estimator, discussed in Hansen et al. (1996) with a quadratic spectral kernel and bandwidth choice according to Andrews (1991). For numerical optimization, the quasi-Newton method proposed simultaneously by Broyden, Fletcher, Goldfarb and Shanno (see Broyden, 1970) is applied.

8.3.1 Data

The empirical analysis covers the period from January 1984 until December 2017. As the analysis is limited to the US futures market, for y_t the WTI benchmark spot price is used as realizations of the oil price and futures prices, traded on the NYMEX for oil price forecasts to the horizons $h = 1, ..., 6$. All prices are deflated by the US CPI.[1]

Formally, one observes in period t the spot price of oil y_t and the futures prices with maturity h, f_{t+h}, interpreted as forecast for period $t+h$. By contrast the forecasts for period t were observed as futures prices in periods $t-h$ for each futures maturity. The forecast error for period t given a futures maturity h is consequently constructed as difference between the observed spot price in period t and futures price in period $t-h$: $e_t = y_t - f_{t-h}$.

It is very important to construct the instrument sets based on the information set $\boldsymbol{w}_t$ available to market participants in period t. In a first step, the instrument sets of the original approach, proposed by Elliott et al. (2005, 2008) are adopted, using for each horizon h, a constant (set A) $\boldsymbol{w}_t = (1)'$. Second, a constant and the lagged forecast error is assumed, corresponding with the horizon in question (set B) and observed in period $t-1$, $\boldsymbol{w}_t = (1, e_{(t-h)-1})'$. Finally, an instrument set is adopted, consisting of the first lag of the real oil price, the first in addition to a constant (set C) $\boldsymbol{w}_t = (1, rpo_{t-1})'$ and an instrument set, consisting of a constant as well as the first and second lags of the real oil price (set D) $\boldsymbol{w}_t = (1, rpo_{t-1}, rpo_{t-2})'$. In both cases the real price of oil is expressed in growth rates.

Finally, a wide range of financial variables is employed in the tradition of Fama and French (1987, 1988) that have been used in the literature to explain returns on commodities futures in general and crude oil futures in particular. The variables employed as well as the corresponding studies are shown in table 8.1. For the sample period covered and given data availability we were able to recreate and include 26 series previously employed in the literature, allowing to construct the following instrument sets $w_t^i = (1, var_{t-1}^i)$ for each variable var_t^i $i = 1, ..., 26$. For the sake of completeness and further robustness, principal component analysis (see e.g. Jolliffe, 2002), is used for aggregating the information into a reduced number of variables. Out of the 26 principal components extracted, only the first 5 components are used that account for 62% of the total information and construct two last instrument sets $w_t^{27} = (1, PC_{t-1}^1, ..., PC_{t-1}^5)$ and the second lagged set $w_t^{28} = (1, PC_{t-2}^1, ..., PC_{t-2}^5)$.

Data sources for the construction of the series are provided in the appendix. The aforementioned studies include a wider range of explanatory variables to explain return on futures prices, some of which are discarded in the analysis because of data availability and non-stationarity in their original.

[1]WTI series are retrieved from the FRED database with the series code CRUDOIL. Futures prices are retrieved from the Thomson Reuters database with the series codes NCL00, NCL02, NCL03, NCL04, NCL05, NCL06. US CPI is retrieved from the FRED database with the series code CPIAUCSL

After the baseline estimation that does not consider structural shifts, we further estimate two specifications that control for possible structural breaks. First, we include the structural break as identified by Hamilton and Wu (2014), modeled as a dummy taking the value of 1 starting in January 2005. Secondly, as a further robustness check we include the shifts estimated by SIS with regard to the real price of oil as presented below. The underlying assumption is that because of the linkages between spot market and futures market for crude oil by an arbitrage condition (Fattouh et al., 2013), structural changes affecting the spot market also might affect the futures market and thus the risk premium, modeled by the forecast error. As a final robustness check, time-varying estimates of the asymmetry parameter α are explored.

8.3.2 Structural breaks with regard to the WTI oil price series

In this section, the R-implementation of SIS (Castle et al., 2015) by Sucarrat et al. (2016) are applied to estimate the mean levels of the WTI spot price of crude oil (expressed in constant 2017 \$) between January 1984 and December 2017. The shifts selected at the 0.1% significance level are shown in figure 8.3.

The first significant shift occurs in July 1991 when the average Dollar value of a barrel of crude oil changes from around \$45 to \$31. A slight reversal to a mean value of \$37 is estimated in January 1999. The level remains stable for five years, but the next two significant shifts are identified in March 2004 and June 2005. A large increase is then estimated in November 2007 leading to a peak in the average value of around around \$120 per barrel, but still below the peak in the real spot price of oil in July 2008 at slightly over \$145 per barrel. An important negative level shift is estimated in November 2008, following the start of the Great Recession. Two positive shifts in May 2009 and in August 2011 mark a recovery of the oil price. In December 2014 however, an important negative shift is identified, stabilizing the expected level around the mark of \$49 per barrel.

Although Hamilton and Wu (2014) specifically limit their analysis to the price of oil futures and find significant changes in the futures market structure from 2005 onwards, a comparison with the estimation results provides acceptable results. With the exception of the first two shift, that slightly compensate each other and are limited in absolute terms, the remaining level shifts estimated by SIS are located around and after the year 2005. Thus, the SIS results and their use as a robustness check in this study seem plausible.

Author	**Predictor variables**	**Instrument set**
Bessembinder (1992)	Returns on CRSP value-weighted equity index	1
	Unexpected CPI inflation	2
	Change in expected CPI inflation	3
	Change in 3-month T-bill rate	4
	Change in the term structure (20YGB – 3-month T-bill)	5
	Change in default premium (Baa – 20YGB)	6
	Unexpected change in U.S. industrial production	7
Bessembinder and Chan (1992)	Dividend yield on CRSP value-weighted equity index	8
De Roon et al. (2000)	Returns on S&P 500 stock price index	9
Sadorsky (2002)	Return on dividend yield on S&P 500 common stock portfolio	10
	Return on junk bond premium (Baa – Aaa)	11
	Return on 3-month T-bill rate	12
	Market portfolio excess return	13
Pagano and Pisani (2009)	Degree of capacity utilization in U.S. manufacturing	14
	Term spreads (2YGB–1YGB, 5YGB–2YGB, 10YGB–5YGB)	15-17
	Composite leading indicator for OECD	18
Hong and Yogo (2012)	Yield spread (Aaa – 1-month T-bill)	19
	Basis by horizon	20-25
Gorton et al. (2012)	Normalized U.S. commercial crude oil inventories (no SPR)	26

Table 8.1: Financial variables employed as instrument sets for the GMM estimation with corresponding sources.

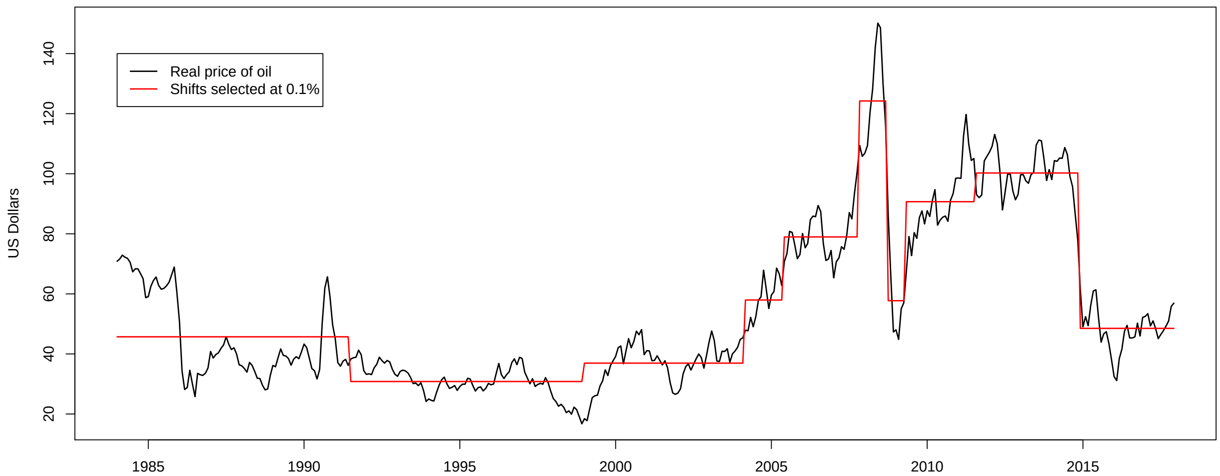

Figure 8.3: Real WTI spot price (in 2017 \$) with steps selected by SIS at 0.1 % over the sample period January 1984 to December 2017 (Source: FRED).

8.3.3 Results under symmetry and asymmetry

In this section the main empirical findings are presented and discussed for each specification:

- The whole sample without taking into account structural breaks,
- the whole sample including the Hamilton and Wu (2014) structural break and
- the structural breaks as implied by SIS.

The results concerning unbiasedness and efficiency tests under symmetry are shown, based on the (modified) MZ regression in equations in equations (8.6) and (8.7). Additionally, the results concerning the estimate of the asymmetry parameter α are presented as well as the rationality (optimality) of the market prediction, if asymmetric preferences are permitted. Finally, time-varying estimates of the asymmetry parameter α resulting from the rolling window estimation are outlined.

The empirical application follows the literature, fixing the curvature parameter to $p = 1$ and $p = 2$ in the EKT loss function (see equation (8.9)) and present the results for both cases. Furthermore, it is worth noting that the results are based on the absolute forecast error terms e_{t+h} for each futures horizon as they point to a better fit under asymmetry. The estimation results expressed in relative forecast errors e_{t+h}/f_{t+h} were comparable when looking at the estimates $\hat{\alpha}$ of the asymmetry parameter, but slightly less consistent across all specifications and also resulted in more frequent optimality rejections under asymmetry.

We recall that the aim of this chapter is to find evidence for a non zero risk premium by applying the EKT framework to crude oil futures and estimate the asymmetry parameter α. Before going into a detailed discussion of the results for each specification, we will briefly point to the main findings that we observe consistently across all specifications and instrument sets.

- Under the assumption of symmetric loss, we are unable to reject unbiasedness and efficiency in the majority of cases using the MZ equations.
- When allowing for asymmetric loss however, we find consistent estimates of the asymmetry parameter $\hat{\alpha} < 0.5$, implying a preference of market participants to under-estimate the crude oil price through futures pricing. This result is in line with the existence of a positive risk premium on crude oil futures.
- The values of $\hat{\alpha}$ have a tendency to decrease with higher futures maturities.
- When fixing $p = 2$ and for the instruments sets containing the individual financial variables (sets 1-26), we estimate inconsistent estimates of α.
- We observe very few rejections of forecast rationality under asymmetric loss. Again, a few exceptions are observed for individual financial instrument sets.

Baseline specification without structural breaks 1984-2017

Table 8.2 shows the results using the basic EKT instrument sets A-D outlined above. The first main column (Results under symmetry) shows the p-values with regard to the null hypothesis of unbiasedness (set A) and efficiency (sets B-D) under symmetric loss and for each maturity $h = 1, ..., 6$. The second main column (Results under asymmetry) contains the estimation result for the asymmetry parameter $\hat{\alpha}$ and the J-statistic with the corresponding p-value for testing forecast optimality (model validity) for the curvature cases $p = 1$ and $p = 2$. As a note, because instrument set A contains a single constant and we estimate a unique parameter $\hat{\alpha}$, the model is exactly identified and the J-statistic is always equal to zero.

In the symmetric case one notes that unbiasedness cannot be rejected for any maturity. The same is mostly true for the null hypothesis of efficiency. Only the efficiency at the 5% significance level can be rejected in the case of the instrument set B for maturities $h = 4$ and $h = 5$. Relaxing the condition of symmetry and looking at the results for $\hat{\alpha}$ that are not rejected by testing the null hypothesis of optimality, a main pattern of the empirical results

		Results under symmetry	**Results under Asymmetry**					
			$p = 1$			$p = 2$		
Horizon	Set	p-val	$\hat{\alpha}$	J-stat	p-val	$\hat{\alpha}$	J-stat	p-val
$h = 1$	A	0.825	0.478	0.000	1.000	0.511	0.000	1.000
	B	0.383	0.479	0.028	0.867	0.474	0.478	0.489
	C	0.322	0.478	0.002	0.967	0.444	1.216	0.270
	D	0.270	0.478	0.205	0.903	0.438	4.507	0.105
$h = 2$	A	0.715	0.458	0.000	1.000	0.518	0.000	1.000
	B	0.876	0.458	0.018	0.894	0.499	0.104	0.747
	C	0.472	0.458	0.189	0.664	0.444	0.691	0.406
	D	0.222	0.459	0.805	0.669	0.445	4.173	0.124
$h = 3$	A	0.753	0.446	0.000	1.000	0.516	0.000	1.000
	B	0.337	0.445	0.056	0.813	0.494	0.256	0.613
	C	0.723	0.447	0.080	0.777	0.455	0.329	0.566
	D	0.395	0.443	1.974	0.373	0.433	3.025	0.220
$h = 4$	A	0.821	0.444	0.000	1.000	0.513	0.000	1.000
	B	**0.042**	0.436	0.728	0.394	0.416	0.855	0.355
	C	0.845	0.445	0.104	0.747	0.465	0.180	0.671
	D	0.123	0.443	5.966	0.051	0.398	4.683	0.096
$h = 5$	A	0.850	0.412	0.000	1.000	0.512	0.000	1.000
	B	**0.020**	0.408	0.896	0.344	0.389	0.422	0.516
	C	0.973	0.411	0.182	0.670	0.521	0.033	0.856
	D	0.271	0.410	0.473	0.790	0.383	2.848	0.241
$h = 6$	A	0.877	0.446	0.000	1.000	0.511	0.000	1.000
	B	0.121	0.453	0.418	0.518	0.378	0.320	0.572
	C	0.946	0.446	0.021	0.884	0.507	0.130	0.719
	D	0.323	0.464	5.119	0.077	0.322	2.453	0.293

Table 8.2: Unbiasedness and efficiency results under symmetry and estimates $\hat{\alpha}$ with the corresponding J-statistic and p-values for the null of optimality under asymmetry for the EKT instruments with no structural breaks.

can be recognized, which are repeatedly estimated throughout this chapter. In the case of $p = 1$, for all futures horizons and all instrument sets, values $\hat{\alpha}$ are estimated that are smaller than 0.5. Furthermore, up to horizon $h = 5$, the value of $\hat{\alpha}$ has a tendency to decrease. It increases slightly for maturity $h = 6$, while remaining below the symmetric value of 0.5. At the 5% significance level one observes no rejections of forecast optimality, implying that the estimates of α are valid.

For $p = 2$, the results show roughly the same pattern with minor differences. With the exception of the cases mentioned below, all estimated values of the symmetry parameter remain $\hat{\alpha} < 0.5$. For the instrument set A (constant) no indication of asymmetry exists, as the point estimates $\hat{\alpha}$ oscillate around 0.5. In the case of the instrument set B (lagged forecast error) for maturities $h = 2$ and $h = 3$ and set C (lagged oil price) for futures maturities $h = 5$ and $h = 6$, comparable values close to 0.5 can be observed. The tendency for the $\hat{\alpha}$ values to decrease with increasing futures maturities is also noted for most sets after $h = 2$. More importantly, no rejections of forecast optimality at the 5% significance level can be detected.

Before looking at the estimates resulting from further instrument sets and model specifications, an exemplary brief interpretation of the $\hat{\alpha}$ value is presented. Let's recall from section 8.2.4 that a value $\hat{\alpha} < 0$ implies a preference to underestimate the spot price of crude oil as the cost of overestimating is higher. For example given $p = 1$, $h = 4$ and instrument set B results in an estimate $\hat{\alpha} = 0.436$, implying $\hat{alpha}/(1 - \hat{\alpha}) = 0.436/0.564 \approx 0.77$ corresponding to an approximate loss difference of 23 per cent. An over-estimation is thus 23 per cent costlier than an under-estimation of the same magnitude. In the case of $p = 1$, $h = 5$ and instrument set B the loss difference corresponds to 31 per cent. Thus on average and over the whole sample, the market has a tendency to underestimate the spot price of crude oil through futures pricing, implying the existence of a positive risk premium on crude oil futures.

The evidence for asymmetry is also confirmed by the estimation using the 26 financial instrument sets. First, as in the previous case, with four exceptions, instrument set 6 (change in default premium Baa-20YGB) for maturity $h = 6$, instrument set 16 (term spread 5YGB-2YGB) for $h = 5$ and instrument sets 15 and 16 (term spreads 2YGB-1YGB and 5YGB-2YGB) for maturity $h = 6$, no rejections of efficiency can be observed using the MZ equations under symmetry. Consequently we primarily focus on the results under asymmetry as shown in figures 8.4 and 8.5 for the curvature cases $p = 1$ and $p = 2$, respectively.

Figure 8.4 shows the GMM estimation results for all 26 financial instrument sets when asym-

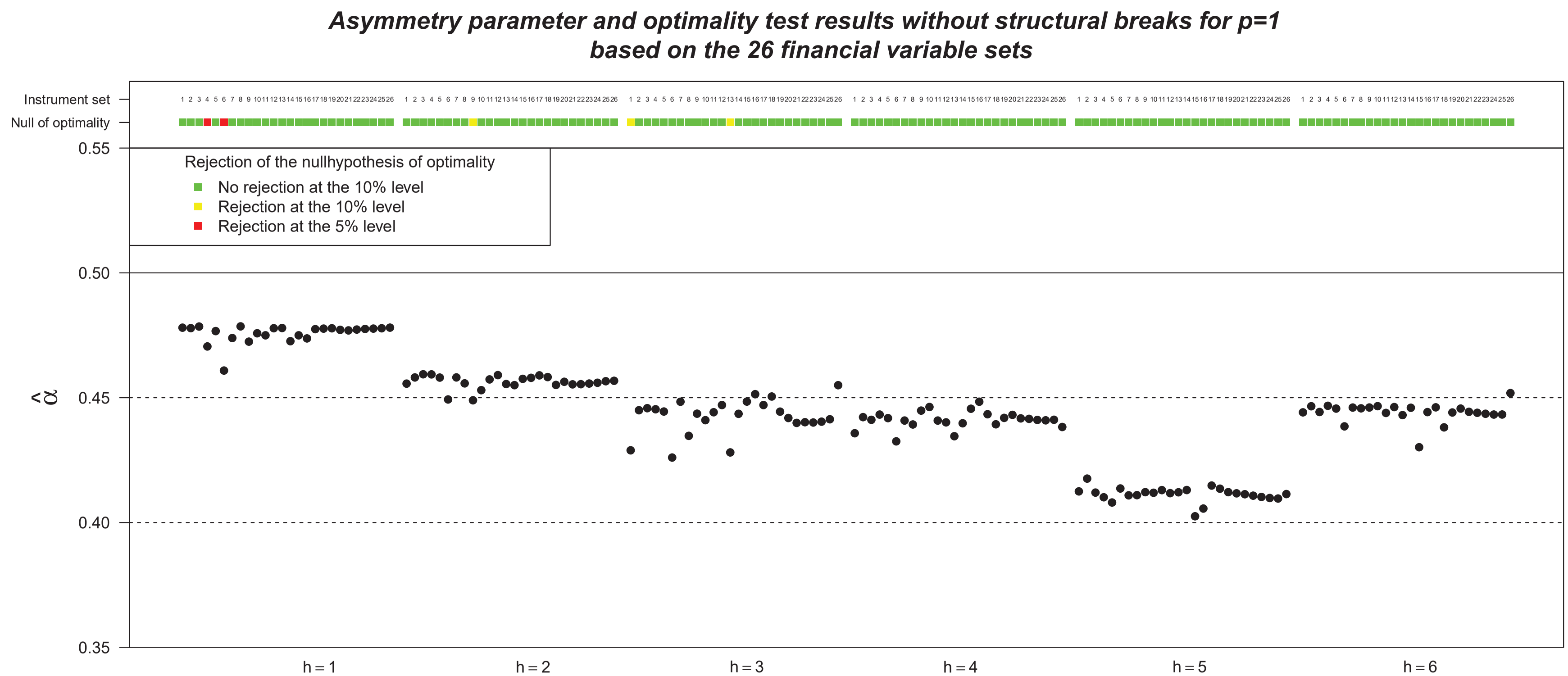

Figure 8.4: GMM estimation results for the 1984-2017 sample without structural breaks. Alpha point estimates while fixing $p = 1$ when considering the 26 financial variable sets.

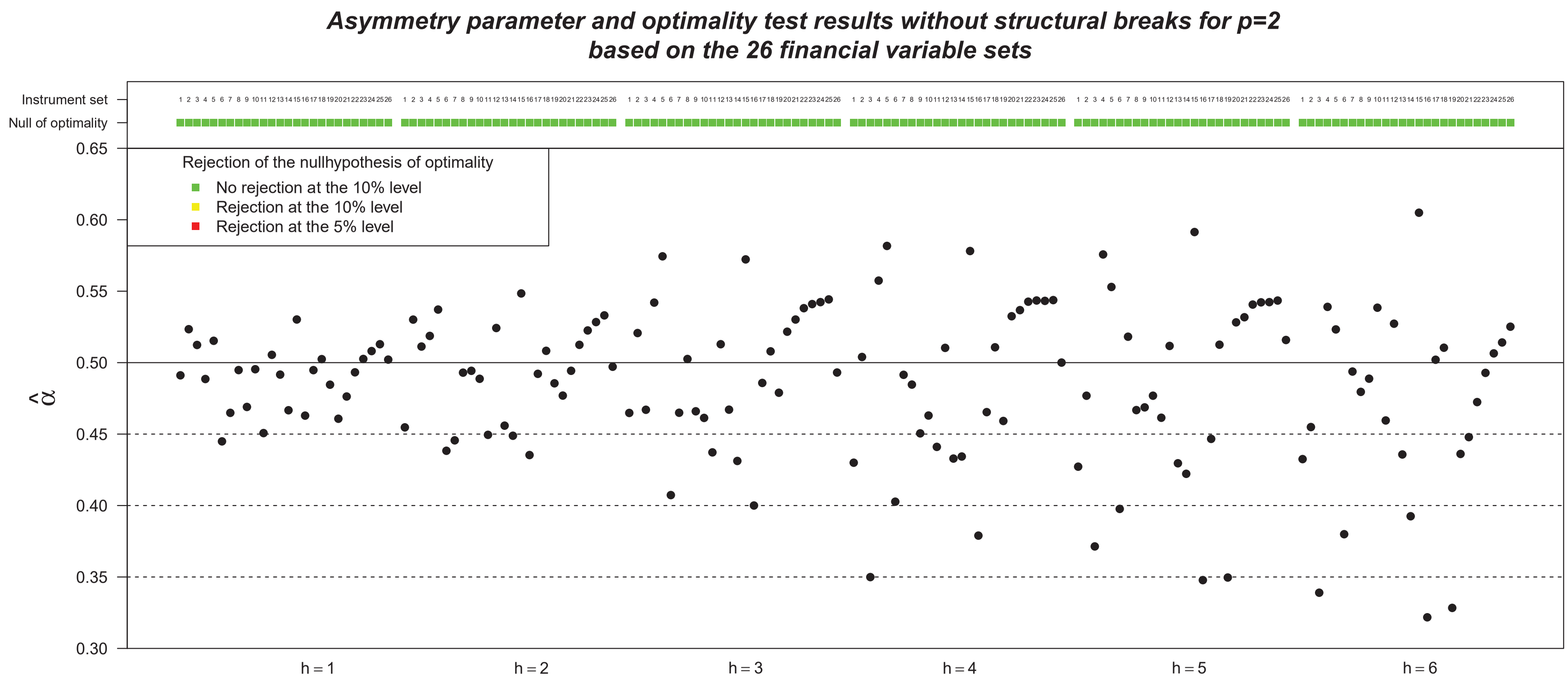

Figure 8.5: GMM estimation results for the 1984-2017 sample without structural breaks. Alpha point estimates while fixing $p = 2$ when considering the 26 financial variable sets.

metry for the case $p = 1$ is allowed. In the upper part of the figure, the rejection of the null hypothesis of forecast optimality (model validity) for all instrument sets across all futures maturities can be seen. Under "green", the null hypothesis of rationality cannot be rejected at the 10% significance level, under "yellow", the null hypothesis can be rejected at the 10% significance level and under "red" the null hypothesis can be rejected at the 5% level.

A small number of optimality rejections becomes evident. We can reject forecast optimality on the 5% significance level for $h = 1$ and instrument sets 4 (change in 3-month T-bill rate) and 6 (change in default premium Baa-20YGB). We can reject forecast optimality on the 10% significance level for $h = 2$ and instrument set 9 (returns on S&P500 index) as well as for $h = 3$ given instrument sets 1 (returns on CRSP value-weighted equity index) and 13 (market portfolio excess return).

The remaining valid point estimates $\hat{\alpha}$ suggest, as the previous results estimated based on the basic EKT instruments, that the market might indeed follow an asymmetric loss function. $\hat{\alpha} < 0.5$ for all cases and we observe a tendency for $\hat{\alpha}$ to get smaller with higher futures maturities. Similarly as before, for $h = 6$ we observe a slight reversal, the estimate $\hat{\alpha}$ however remains far below the symmetric case of 0.5.

Figure 8.5 repeats the results for the case $p = 2$. Surprisingly, given that we observe no rejection in forecast optimality, whereas in the previous case all instrument sets resulted in similar point estimates $\hat{\alpha}$, no such clustering for $p = 2$ can be identified. Some tendency for instrument sets to result in smaller $\hat{\alpha}$ values for longer maturities, for other instrument sets no real change can be observed. In some cases, values above 0.5 can be seen consistently and no rejections of optimality are registered. The results for $p = 2$ are surprising, especially when keeping in mind the similarity of the results for $p = 1$ and $p = 2$ when using the basic EKT instrument sets from table 8.2. Two explanations come to mind. First, Baumeister and Kilian (2016a) observe partially substantial differences when evaluating the OLS-estimates of the risk premium based on equation (8.2). This would indicate the presence of weak instruments in some instrument sets. Second, as will be confirmed by the results under the specification with the Hamilton & Wu (HW) structural break and the SIS structural breaks, the inconsistency is observed uniquely for $p = 2$, pointing to the possibility that the shape of the loss function is simply better represented by $p = 1$.

Finally, table 8.3 contains the results when the two instrument sets based on principal components (sets 27 and 28) are taken into consideration. The results are consistent and comparable

		Results under symmetry	Results under Asymmetry					
			$p = 1$			$p = 2$		
Horizon	Set	p-val	$\hat{\alpha}$	J-stat	p-val	$\hat{\alpha}$	J-stat	p-val
$h = 1$	27	0.851	0.466	4.076	0.539	0.462	1.839	0.871
	28	0.639	0.469	4.594	0.467	0.509	3.348	0.647
$h = 2$	27	0.848	0.450	1.830	0.872	0.468	2.497	0.777
	28	0.638	0.448	5.264	0.384	0.511	4.113	0.533
$h = 3$	27	0.767	0.428	3.774	0.582	0.405	4.250	0.514
	28	0.643	0.449	2.693	0.747	0.527	3.432	0.634
$h = 4$	27	0.722	0.439	2.424	0.788	0.399	4.194	0.522
	28	0.571	0.444	1.782	0.878	0.455	2.290	0.808
$h = 5$	27	0.744	0.412	1.929	0.859	0.422	2.177	0.824
	28	0.408	0.406	1.504	0.913	0.392	2.308	0.805
$h = 6$	27	0.711	0.448	2.921	0.712	0.402	1.583	0.903
	28	0.663	0.443	1.198	0.945	0.368	1.779	0.879

Table 8.3: Efficiency results under symmetry and estimates $\hat{\alpha}$ with the corresponding J-statistic and p-values for the null of optimality under asymmetry for the instrument sets based on principal components with no structural breaks.

with those implied by the EKT instruments, as seen in table 8.2 above. In no case, can efficiency be rejected under symmetric loss. In the case of $p = 1$, all values of $\hat{\alpha}$ are smaller than 0.5, showing a slight tendency in decreasing values for increasing futures maturities, with the exception for $h = 6$. For $p = 2$ very similar results are estimated. With the three exceptions regarding instrument set 28 for maturities $h = 1$, $h = 2$ and $h = 3$, values $\hat{\alpha}$ smaller than 0.5 are estimated, with a tendency to decrease with longer futures horizons. For both cases $p = 1$ and $p = 2$ optimality is never rejected, implying valid results for the estimates $\hat{\alpha}$. As for the previous instrument sets, over the majority of horizons we estimate values of $\hat{\alpha} < 0.5$ that are rational with a loss function that penalizes over-predictions in comparison to under-predictions, as we expect given the existence of a positive risk premium.

		Results under symmetry	Results under Asymmetry					
			$p = 1$			$p = 2$		
Horizon	Set	p-val	$\hat{\alpha}$	J-stat	p-val	$\hat{\alpha}$	J-stat	p-val
$h = 1$	A	0.730	0.478	0.000	1.000	0.511	0.000	1.000
	B	0.408	0.479	0.176	0.916	0.476	0.504	0.777
	C	0.393	0.478	0.173	0.917	0.451	1.362	0.506
	D	0.298	0.479	0.347	0.951	0.438	4.714	0.194
$h = 2$	A	0.517	0.458	0.000	1.000	0.518	0.000	1.000
	B	0.749	0.459	0.194	0.907	0.494	0.303	0.859
	C	0.463	0.459	0.390	0.823	0.446	0.637	0.727
	D	0.230	0.460	1.097	0.778	0.433	4.420	0.220
$h = 3$	A	0.386	0.446	0.000	1.000	0.516	0.000	1.000
	B	0.175	0.445	0.065	0.968	0.470	0.466	0.792
	C	0.512	0.447	0.079	0.961	0.458	0.312	0.856
	D	0.283	0.443	2.009	0.570	0.424	3.194	0.363
$h = 4$	A	0.324	0.444	0.000	1.000	0.513	0.000	1.000
	B	**0.016**	0.435	0.785	0.675	0.415	0.938	0.626
	C	0.532	0.444	0.177	0.915	0.466	0.273	0.872
	D	0.163	0.448	5.769	0.123	0.395	4.695	0.196
$h = 5$	A	0.361	0.412	0.000	1.000	0.512	0.000	1.000
	B	**0.004**	0.409	0.952	0.621	0.395	0.439	0.803
	C	0.640	0.412	0.256	0.880	0.498	0.421	0.810
	D	0.232	0.411	0.580	0.901	0.378	2.854	0.415
$h = 6$	A	0.386	0.446	0.000	1.000	0.511	0.000	1.000
	B	**0.037**	0.458	0.696	0.706	0.390	0.341	0.843
	C	0.612	0.451	0.170	0.918	0.474	0.369	0.831
	D	0.233	0.466	5.390	0.145	0.335	2.541	0.468

Table 8.4: Unbiasedness and efficiency results under symmetry and estimates $\hat{\alpha}$ with the corresponding J-statistic and p-values for the null of optimality under asymmetry for the EKT instruments with the Hamilton and Wu (2014) structural break.

Specification with the Hamilton and Wu (2014) structural break 1984-2017

In the following, the results under symmetric and asymmetric loss, controlling for the Hamilton and Wu (2014) structural break in 2005 are discussed. Focus is laid on the similarities and the few differences in comparison to the specification without structural breaks, discussed in the previous subsection. Table 8.4 shows the results using the basic EKT instrument sets, while controlling for the Hamilton and Wu (2014) structural break. The evaluation of unbiasedness under symmetric loss, the null hypothesis for any futures maturity (even though the p-values are smaller) can be rejected in no case. Efficiency is rejected for instrument set B for maturities $h = 4$, $h = 5$ and $h = 6$, thus giving a weak signal of possible asymmetry. Relaxing the condition of symmetric loss and looking at the results for $p = 1$, leads to the same pattern as before: All estimates $\hat{\alpha}$ of the asymmetry parameter are smaller than 0.5 and a consistent decrease of all estimates the higher the maturity with the exception for $h = 6$ (still lower than 0.5) can be identified. The results for $p = 2$ are very similar: With the exception of the instrument set A (constant), for which we estimate values $\hat{\alpha}$ very slightly above 0.5 for all maturities, all other estimates are smaller than 0.5 and decrease with increasing futures maturities. For both cases $p = 1$ and $p = 2$ optimality is never rejected implying valid results for the estimates $\hat{\alpha}$ that confirm the previous pattern.

Figures 8.6 and 8.7 show the results of the estimations under asymmetry with regard to the 26 financial instrument sets. As before, using the MZ framework, very few rejections of efficiency in the symmetric case can be observed. Thus we focus on the estimates under asymmetry. In figure 8.6 optimality is rejected on the 10% significance level for a single case, set 6 (change in default premium Baa-20YGB) for $h = 1$. For all remaining valid point estimates for the asymmetry parameter $\hat{\alpha}$, one notes that $\hat{\alpha} < 0.5$ for all cases, and a tendency for $\hat{\alpha}$ to get smaller with higher futures maturities. Again, for $h = 6$, a slight reversal is noticeable, however, far below the symmetric case of 0.5.

The case $p = 2$ similarly diverges from the previous results in the sense that no such consistent clustering for all 26 instrument sets is recognizable. In comparison with the specification that does not control for structural breaks, a much higher number of values $\hat{\alpha}$ that are smaller than 0.5 are observed. The decreasing pattern with increasing futures maturities is also more visible, in comparison with the specification without structural breaks. Optimality cannot be rejected for any case.

Finally, table 8.5 contains the results for the two instrument sets based on principal components, extracted from the 26 financial variables (sets 27 and 28). Again, the results are consistent and comparable with the previously mentioned patterns. In no case can efficiency be rejected under symmetric loss. In the case of $p = 1$, all values of $\hat{\alpha}$ are smaller than 0.5 and show a slight tendency to decrease for increasing maturities. The result for $h = 6$ represents the exception of the overall trend, whereas for $p = 2$ very similar results to $p = 1$ are estimated.

Besides the same three exceptions regarding instrument set 28 for futures maturities $h = 1$, $h = 2$ and $h = 3$, consistently values $\hat{\alpha}$ smaller than 0.5 are reported, with a tendency to decrease with longer futures horizons. For both cases $p = 1$ and $p = 2$ optimality is never

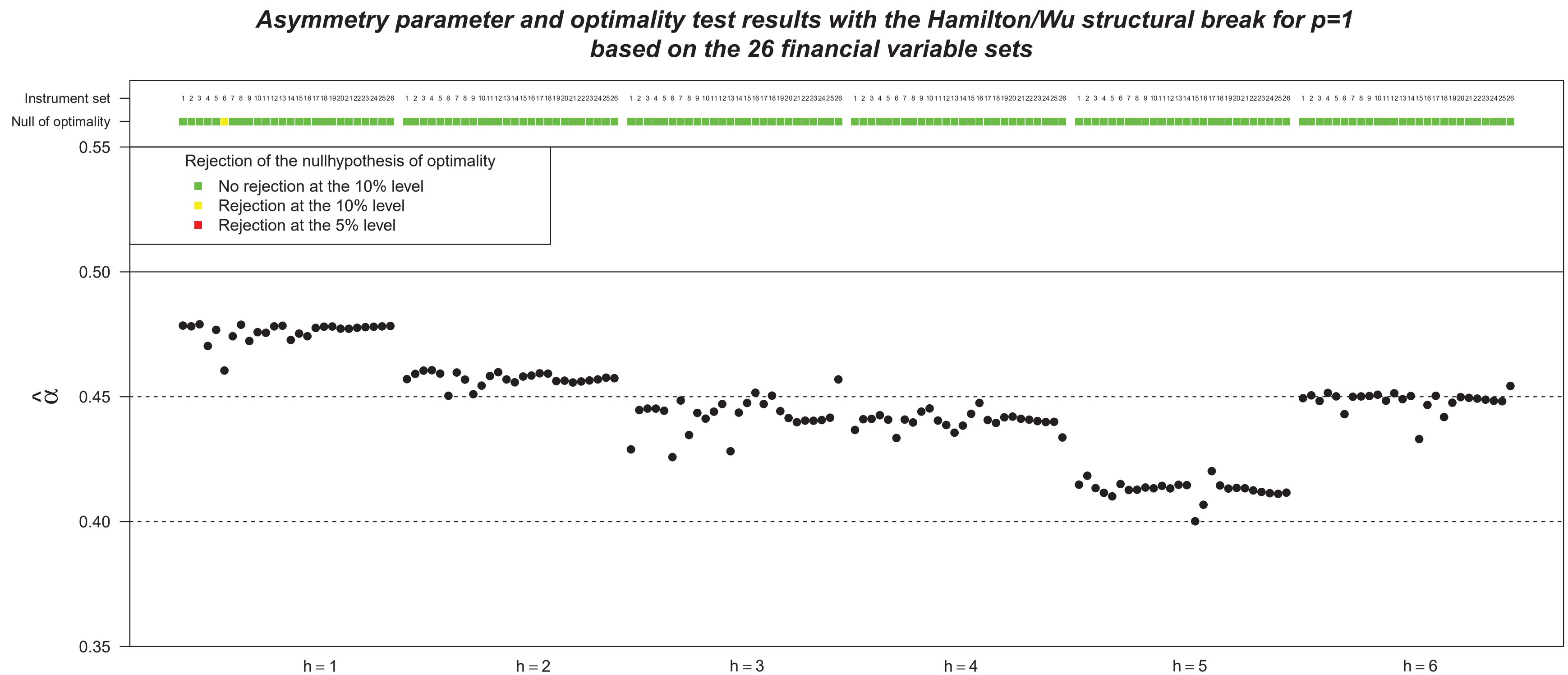

Figure 8.6: GMM estimation results for the 1984-2017 sample with the Hamilton and Wu (2014) structural break. Alpha point estimates while fixing $p = 1$ when considering the 26 financial variable sets.

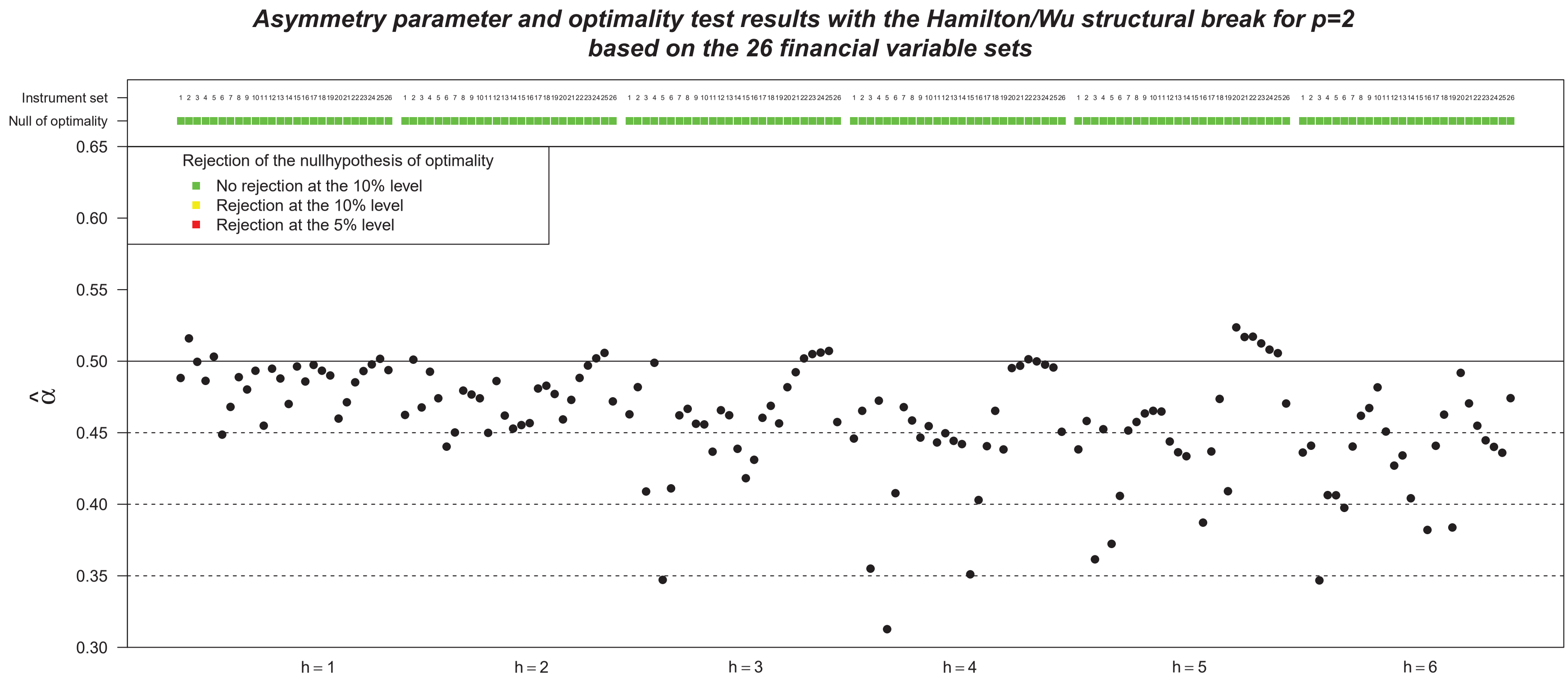

Figure 8.7: GMM estimation results for the 1984-2017 sample with the Hamilton and Wu (2014) structural break. Alpha point estimates while fixing $p = 2$ when considering the 26 financial variable sets.

		Results under symmetry	Results under Asymmetry					
			$p = 1$			$p = 2$		
Horizon	Set	p-val	$\hat{\alpha}$	J-stat	p-val	$\hat{\alpha}$	J-stat	p-val
$h = 1$	27	0.871	0.465	4.624	0.593	0.463	1.949	0.924
	28	0.679	0.469	5.238	0.514	0.514	3.306	0.770
$h = 2$	27	0.786	0.451	1.902	0.929	0.472	2.508	0.868
	28	0.654	0.450	5.320	0.504	0.522	4.131	0.659
$h = 3$	27	0.658	0.429	3.897	0.691	0.385	4.371	0.627
	28	0.580	0.450	2.699	0.846	0.548	3.526	0.741
$h = 4$	27	0.678	0.440	2.515	0.867	0.368	4.467	0.614
	28	0.453	0.444	1.824	0.935	0.457	2.542	0.864
$h = 5$	27	0.674	0.415	2.123	0.908	0.432	2.286	0.892
	28	0.390	0.408	1.618	0.951	0.377	2.394	0.880
$h = 6$	27	0.582	0.455	3.284	0.773	0.414	1.565	0.955
	28	0.593	0.446	1.458	0.962	0.367	1.790	0.938

Table 8.5: Efficiency results under symmetry and estimates $\hat{\alpha}$ with the corresponding J-statistic and p-values for the null of optimality under asymmetry for the instrument sets based on principal components with the Hamilton and Wu (2014) structural break.

rejected implying valid results for the estimates $\hat{\alpha}$.

The estimates when controlling for the HW structural break thus confirm the previous baseline estimates. First, the MZ-equations do not reject uniasedness and only in some cases reject efficiency. Values $\hat{\alpha} < 0.5$ are consistently estimated based on all instrument sets, implying that market participants have a preference to under-estimate the sport price of oil through futures pricing. This is in line with the existence of a positive risk premium on futures contracts. We observe the same pattern for $\hat{\alpha}$ to decrease with higher maturities, indicating increasing risk premia for longer horizons. The inconsistent results based on the 26 financial instrument are again observed. The possibility of weak instruments or a better fit when fixing $p = 1$ for the loss function are again confirmed by this second specification.

		Results under symmetry	Results under Asymmetry					
			$p=1$			$p=2$		
Horizon	Set	p-val	$\hat{\alpha}$	J-stat	p-val	$\hat{\alpha}$	J-stat	p-val
$h=1$	A	0.767	0.478	0.000	1.000	0.511	0.000	1.000
	B	0.615	0.476	3.976	0.948	0.457	5.807	0.831
	C	0.525	0.472	3.726	0.959	0.442	7.698	0.658
	D	0.291	0.471	3.859	0.974	0.424	11.404	0.410
$h=2$	A	0.759	0.458	0.000	1.000	0.518	0.000	1.000
	B	0.973	0.462	3.800	0.956	0.507	7.511	0.676
	C	0.711	0.462	3.628	0.963	0.483	8.321	0.598
	D	0.284	0.459	5.331	0.914	0.447	12.792	0.307
$h=3$	A	0.822	0.446	0.000	1.000	0.516	0.000	1.000
	B	0.243	0.380	7.687	0.659	0.322	11.906	0.291
	C	0.892	0.404	5.215	0.876	0.492	7.624	0.665
	D	0.368	0.380	7.872	0.725	0.406	13.147	0.284
$h=4$	A	0.855	0.444	0.000	1.000	0.513	0.000	1.000
	B	**0.038**	0.376	14.252	0.162	0.212	11.247	0.339
	C	0.956	0.444	10.726	0.379	0.555	5.918	0.822
	D	0.168	0.427	17.533	**0.093**	0.412	14.257	0.219
$h=5$	A	0.724	0.412	0.000	1.000	0.512	0.000	1.000
	B	**0.011**	0.440	14.318	0.159	0.153	11.070	0.352
	C	0.867	0.426	10.608	0.389	0.553	9.357	0.499
	D	0.192	0.424	11.644	0.391	0.433	11.398	0.411
$h=6$	A	0.620	0.446	0.000	1.000	0.511	0.000	1.000
	B	**0.032**	0.548	7.685	0.660	0.175	9.894	0.450
	C	0.776	0.512	5.839	0.829	0.556	11.551	0.316
	D	0.162	0.560	15.670	0.154	0.477	12.225	0.347

Table 8.6: Unbiasedness and efficiency results under symmetry and estimates $\hat{\alpha}$ with the corresponding J-statistic and p-values for the null of optimality under asymmetry for the EKT instruments with structural breaks implied by SIS.

Specification with the SIS structural breaks 1984-2017

In the following, the results under symmetric and asymmetric loss are highlighted under conditions of controlling for structural breaks, as estimated by SIS (see section 8.3.2). Table 8.6 shows the results under symmetry and asymmetry based on the baseline instrument sets proposed by EKT. With a few exceptions, the main patterns previously mentioned are confirmed. In no case can unbiasedness be rejected. Efficiency on the other hand is rejected when considering instrument set B for maturities $h=4$, $h=5$ and $h=6$.

Under asymmetry while fixing $p=1$, all estimates $\hat{\alpha}$ are smaller than 0.5, with the exception

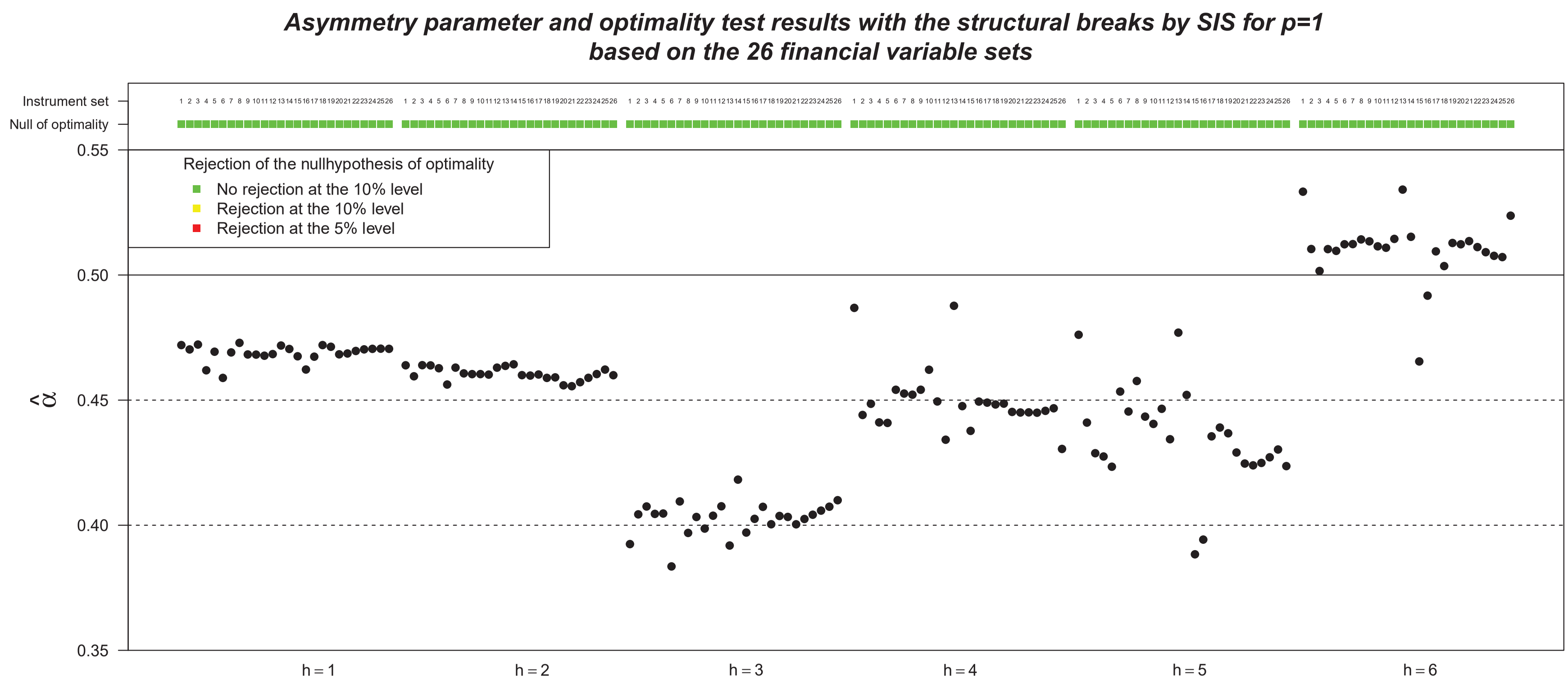

Figure 8.8: GMM estimation results for the 1984-2017 sample with the SIS structural breaks. Alpha point estimates while fixing $p = 1$ when considering the 26 financial variable sets.

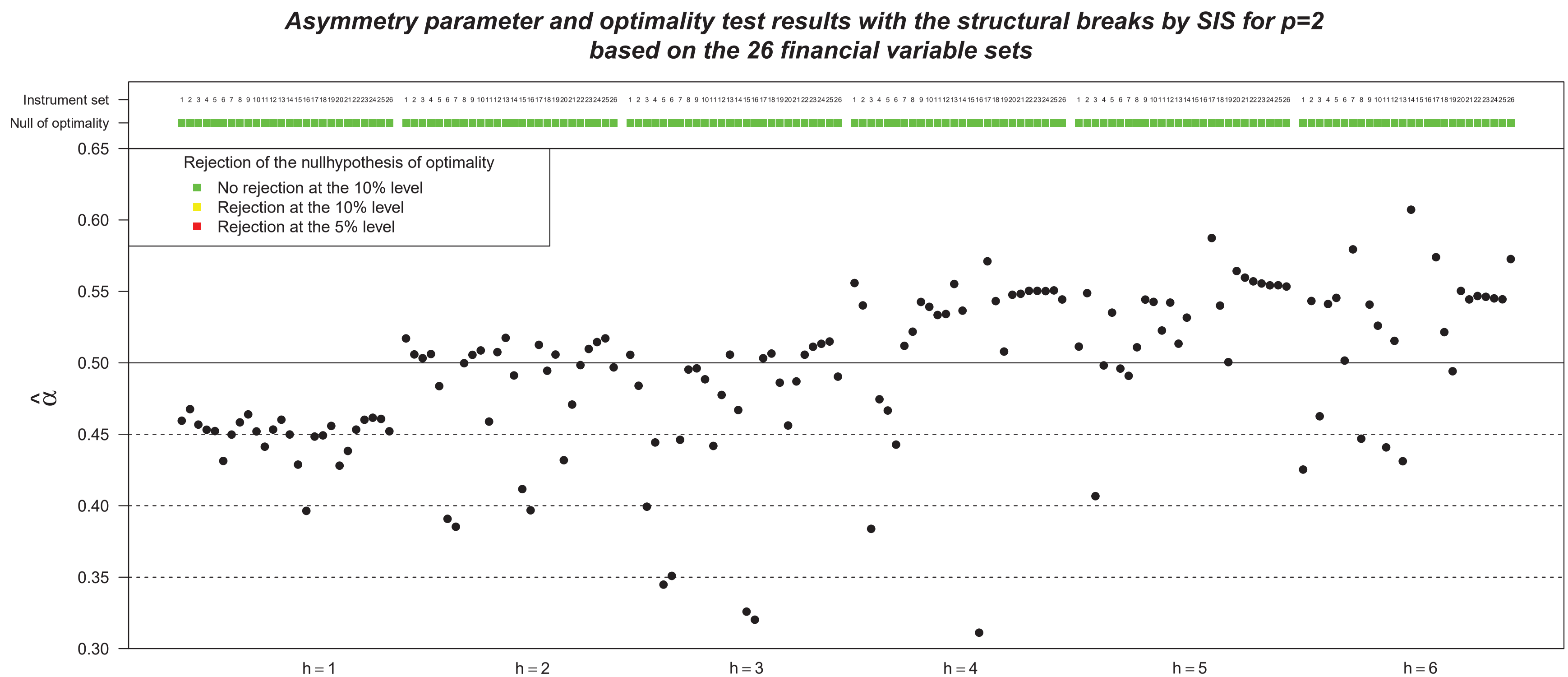

Figure 8.9: GMM estimation results for the 1984-2017 sample with the SIS structural breaks. Alpha point estimates while fixing $p = 2$ when considering the 26 financial variable sets.

of sets B and D for $h = 6$. A consistent decrease of all estimates up to $h = 3$ can be observed, where $\hat{\alpha}$ stabilizes for subsequent maturities. The results for $p = 2$ point to the same direction. With the exception on the subsequent mentioned cases, all estimates $\hat{\alpha} < 0.5$ with a tendency to decrease up to horizon $h = 3$. Values $\hat{\alpha} > 0.5$ appear for all estimates based on set A, set B for $h = 2$ and set C for $h = 4$, $h = 5$ and $h = 6$. Finally, optimality is only rejected in the case of $p = 1$ for set D and $h = 4$, implying that valid results for the estimates $\hat{\alpha}$.

Figures 8.8 and 8.9 show the results of the estimation under asymmetry for the 26 financial instrument sets. Comparable with previous results, very few rejections of efficiency in the symmetric case can be observed so that again, we focus on the estimates under asymmetry. As seen in figure 8.8, optimality cannot be rejected for any estimate. With respect to the point estimates for the asymmetry parameter $\hat{\alpha}$, the conclusion can be drawn that $\hat{\alpha} < 0.5$ for all cases and that there is a tendency for $\hat{\alpha}$ to get smaller up to $h = 3$. A reversal for $h = 4$ and $h = 5$ can be seen, however, still remaining below 0.5. For $h = 6$, on the other hand, one even observes significantly higher values than 0.5 for $\hat{\alpha}$.

The case $p = 2$ differs significantly from previous results with respect to a reversal of the pattern. Whereas estimates for $h = 1$ are compatible with previous values $\hat{\alpha} < 0.5$, an increase in the estimates $\hat{\alpha}$ can be observed, starting with a maturity $h = 2$ up to $h = 6$ which are consistently above 0.5. Optimality is again not rejected for any case. This results seem to suggest that market participants prefer an over-estimation of the spot price of crude oil when pricing crude oil futures, implying that the risk premium is negative.

Finally, table 8.7 contains the results, taking into consideration the two instrument sets based on principal components, extracted from the 26 financial variables (sets 27 and 28). The results are again comparable with the patterns previously highlighted. In no case can efficiency under symmetric loss be rejected. Concerning the case for $p = 1$, all values of $\hat{\alpha}$ are smaller than 0.5 with the exception of set 27 for $h = 6$. A slight tendency in decreasing values $\hat{\alpha}$ for increasing futures maturities up to $h = 3$ emerges. For $p = 2$ very similar results are shown. Besides the exceptions for set 28 and $h = 4$ as well as for set 27 and $h = 6$, consistently values $\hat{\alpha}$ smaller than 0.5 appear, with slight tendencies to decrease up to the futures horizon $h = 3$. For both cases $p = 1$ and $p = 2$ optimality is never rejected, implying valid results of the estimates $\hat{\alpha}$.

The estimates when controlling for the SIS structural breaks thus also mainly confirm the previous baseline estimates. As in the HW case, the MZ-equations do not reject uniasedness

		Results under symmetry	Results under Asymmetry					
			$p = 1$			$p = 2$		
Horizon	Set	p-val	$\hat{\alpha}$	J-stat	p-val	$\hat{\alpha}$	J-stat	p-val
h=1	27	0.826	0.460	11.465	0.649	0.432	8.420	0.866
	28	0.425	0.466	10.942	0.691	0.486	16.037	0.311
h=2	27	0.702	0.462	6.912	0.938	0.471	9.999	0.762
	28	0.639	0.456	8.995	0.831	0.481	12.756	0.546
h=3	27	0.596	0.376	9.123	0.823	0.336	11.588	0.639
	28	0.621	0.396	8.108	0.884	0.468	13.232	0.508
h=4	27	0.328	0.479	12.377	0.576	0.337	10.259	0.743
	28	0.501	0.452	11.546	0.643	0.512	7.878	0.896
h=5	27	0.226	0.444	12.017	0.605	0.402	10.189	0.748
	28	0.410	0.419	10.467	0.727	0.482	7.968	0.891
h=6	27	0.260	0.528	10.333	0.737	0.603	13.531	0.485
	28	0.622	0.482	7.203	0.927	0.463	8.482	0.863

Table 8.7: Efficiency results under symmetry and estimates $\hat{\alpha}$ with the corresponding J-statistic and p-values for the null of optimality under asymmetry for the instrument sets based on principal components with the SIS structural breaks.

and only in some cases reject efficiency. Values $\hat{\alpha} < 0.5$ are consistently estimated based on all instrument sets, implying that market participants have a preference to under-estimate the sport price of oil through futures pricing. This is in line with the existence of a positive risk premium on futures contracts. We observe the same pattern for $\hat{\alpha}$ to decrease with higher maturities, indicating increasing risk premia for longer horizons. The inconsistent results based on the 26 financial instrument are again observed when fixing $p = 2$. Moreover the results even contradict all previous estimates as we observe a much larger number of estimates $\hat{\alpha} > 0.5$. The possibility of weak instruments or a better fit when fixing $p = 1$ are again confirmed.

Rolling window estimates

In the following, the rolling-window estimates $\hat{\alpha}$ using the basic EKT instrument set B are presented. The results rely on a window of 72 months, the first value for $\hat{\alpha}$ is estimated for December 1989, using the window January 1984 to December 1989, the last value for December 2017 based on the window January 2012 to December 2017.

An important peculiarity of the first rolling-window results, is the sensitivity of the EKT framework against outliers, when the estimation is based upon shorter sample sizes in comparison to the whole sample. This materializes by highly volatile and extreme values of $\hat{\alpha}$, in combination with a very high number of optimality rejections within each window.

Figure 8.10 provides an example of the evolution of $\hat{\alpha}$ for futures maturity $h = 3$, using instrument set B (lagged forecast error) and setting $p = 1$. Not only are estimates $\hat{\alpha}$ closer to 1 or 0 more difficult to interpret, but jumps from one extreme to another are hardly justifiable, keeping in mind, that differences in the sample originate from the addition and removal of a single observation only by advancing the sample window. In the particular case shown in figure 8.10 we even see that optimization in the GMM estimation doesn't result in convergence November 2014, as can be seen in a negative estimate $\hat{\alpha}$. Such sensitivity to outliers can be corrected by applying IIS, as in the previous chapter. We thus apply IIS to the forecast error series for each futures maturity and each estimation window and set the significance level of variable retention to 0.1%. On average 1.8 outliers were detected per estimation window. Again, the estimation is executed while fixing $p = 1$ and $p = 2$. All results confirm and complement the patterns that we observed previously over the three specifications covering the whole sample.

The results when controlling for outliers based on instrument set B can be seen in figure 8.11. Irrespective of $p = 1$ or $p = 2$ and the horizon h, the estimated asymmetry parameter $\hat{\alpha}$ oscillates around the value of 0.5 during the 1990s. It implies no clear market preference to under-estimate or over-estimate the spot price of crude oil by future prices. In other words, no clear indication of a non-zero risk premium can be observed. After a consistent temporary increase around 1999 across all futures horizons, overall values $\hat{\alpha}$ are observed which decrease from 2000 onwards and remain consistently below the value of 0.5. $\hat{\alpha}$ return to the symmetric case in 2016 for $p = 1$. For $p = 2$ the value for $\hat{\alpha}$ returns to 0.5 in late 2009. Within this time frame, futures prices thus reflect clear tendencies of market participants to highly under-estimate the spot price of crude oil, in fixing futures prices. This is consistent with a positive risk premium on crude oil futures. We recall that it is within this period, that the

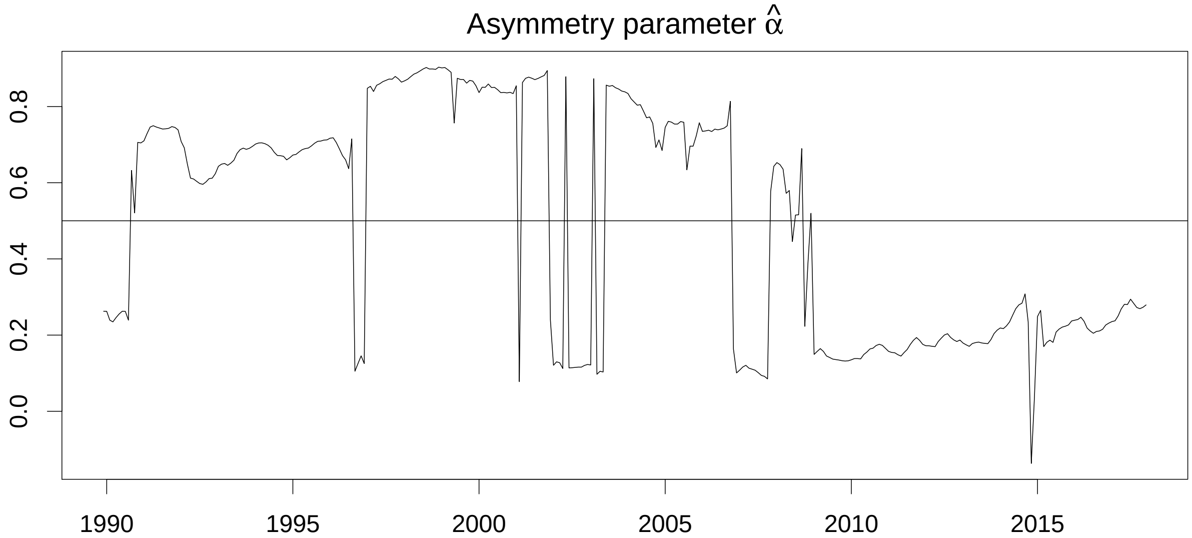

Figure 8.10: Rolling window estimate of alpha without controlling for outliers (length=72 months, p=1, h=3, instrument set B).

spot market experienced very strong prices increases accompanied by a high volatility.

The end of the analyzed sample period is characterized by values of $\hat{\alpha} > 0.5$, especially for $p = 2$. So contrary to previous episodes, market participants over-estimated the spot price of oil through futures pricing, resulting in negative risk premia on crude oil futures. In other words, the buyers of futures contracts pays the premium to the short side of the contract. Such observation seem consistent with the increasing financialization of futures markets as observed by Hamilton and Wu (2014). The main difference between the results for $p = 1$ and $p = 2$ are reflected in a stronger preference to over-estimate the spot prices, when $p = 2$. Furthermore we observe again, a clear tendency that higher futures maturities are associated with values of $\hat{\alpha}$ which are further away from the symmetric value of 0.5. Here again, we relate the results to the developments on the spot market. Since 2015 the oil price experienced much less volatile movements and remained stable around a value of \$50 per barrel.

As an important complementing note, the sample was also tested for unbiasedness and inefficiency using the MZ-framework in every window. The results (not shown in detail) complement the estimates of $\hat{\alpha}$ as would be expected. Within those windows that result estimates for $\hat{\alpha}$ which strongly deviate from the symmetric value of 0.5 unbiasedness and efficiency based on instrument set B can be consistently rejected, with higher rejection rates for increasing futures maturities.

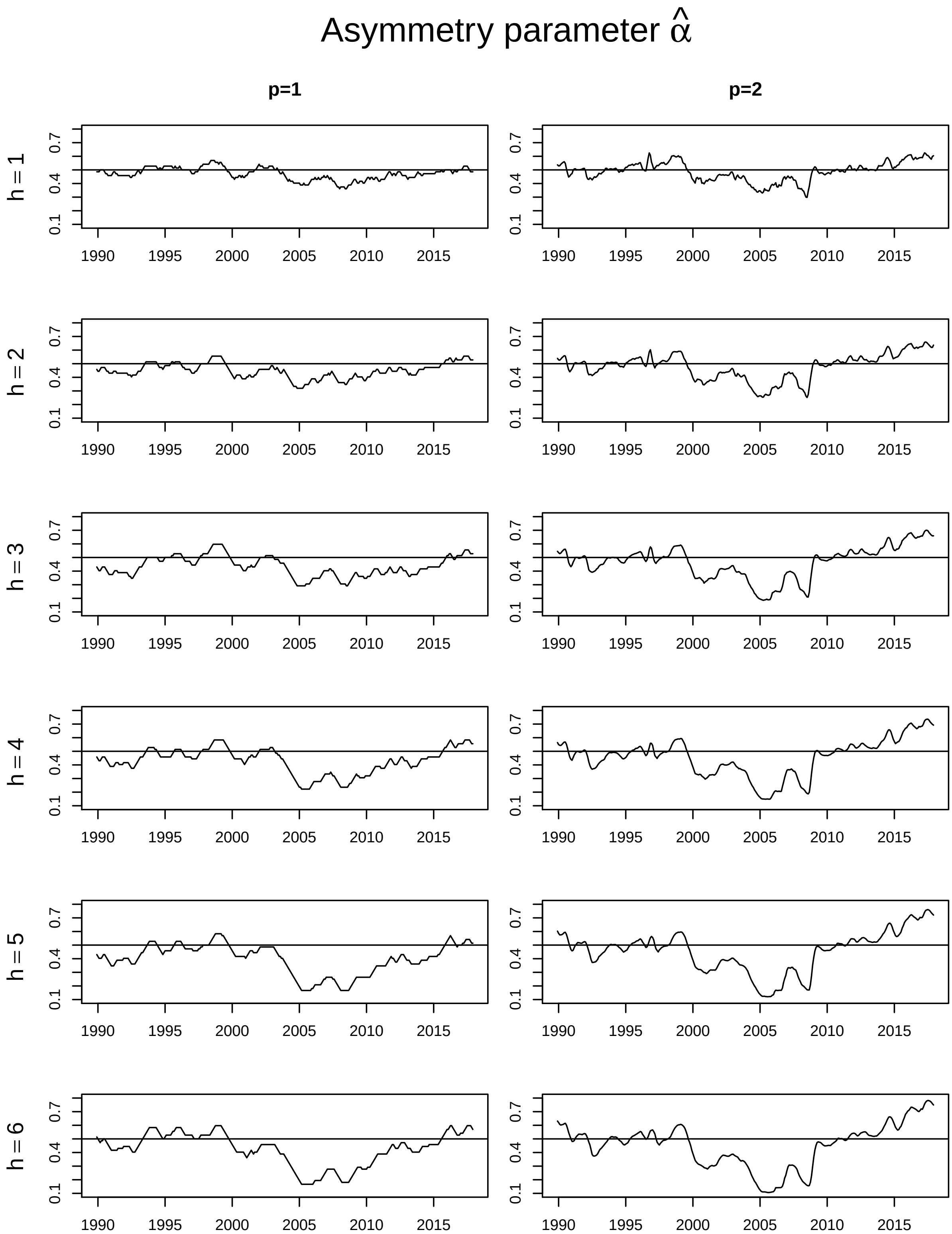

Figure 8.11: Rolling window estimate of alpha without with IIS December 1989 to December 2017 (length=72 months, instrument set B).

8.4 Summary and discussion

This chapter used the framework of Elliott et al. (2005, 2008) to evaluate whether oil market forecasts follow an asymmetric loss function when looking at crude oil futures. Asymmetric loss implies that a positive forecast error is not associated with the same loss or cost as a negative forecast error of the same magnitude. The theoretical justification in assuming the existence of an asymmetric loss when market participants fix futures prices is the presence of a risk premium different from zero. It is common empirical knowledge that crude oil futures - although fixing the quantity and price of crude oil to be delivered in a future period - perform poorly, when they were used as predictors of the spot price of crude oil in a future period.

Using a monthly sample between January 1984 and December 2017, first unbiasedness and efficiency were evaluated under symmetric loss criteria. All tests, even when controlling for the structural break identified by Hamilton and Wu (2014) or the structural breaks estimated by SIS, could not reject unbiasedness. Similarly efficiency could not be rejected in the overwhelming number of cases across all instrument sets and specifications. The classical MZ-framework, even when controlling for structural breaks, thus provided no indication of asymmetric loss over the whole sample.

The latter is relevant in view of the results achieved in a second step under asymmetric loss. Previous individual and institutional forecast evaluation studies (see e.g. Romer and Romer, 2000; Sims, 2002; Capistrán, 2008; Sinclair et al., 2010; Patton and Timmermann, 2012; Rossi and Sekhposyan, 2016) which rely on the use of quarterly data over shorter samples, already needed a modification of the MZ-framework for the evaluation of unbiasedness and inefficiency under symmetric loss by controlling for structural breaks. We recall that the rejection of unbiasedness or efficiency under symmetric loss is a first indication for the possibility of an asymmetric loss function characterizing the forecast Behavior. Our analysis suggests that in order to accommodate data at higher frequency and higher volatility the framework requires further modifications. Recall, that the oil market is characterized by the occurrence of supply and demand shocks that make the data very noisy. As seen in the rolling window estimation, by controlling for outliers with IIS within smaller sample lengths, the expected rejections of unbiasedness and efficiency under symmetry is observed. This in accordance with highly asymmetric values of $\hat{\alpha}$, which were also estimated within those same windows.

Asymmetry, detected on the basis of GMM estimation of the flexible loss function, proposed by Elliott et al. (2005, 2008), is highly robust against different specifications and a large number of instrument sets. Forecast optimality under asymmetric loss is confirmed resulting implying following main pattern: Over the whole sample - with and without controls for structural breaks - the results confirm the presence of a positive risk premium on crude oil futures under the use of the original EKT instrument sets. All estimates of the asymmetry parameter α were smaller than the symmetric value of 0.5, meaning that crude oil futures underestimate the spot price of crude oil consistently. A further conclusion consistent with the theory of the risk premium was that for higher maturities the value of $\hat{\alpha}$ decreased. In other words, for higher horizons the risk premium implied by the under-estimation was more important.

Under asymmetry, estimates on the basis of the 26 financial variable instrument sets basically confirm the results, implied by the EKT instruments. In most cases, values of $\hat{\alpha}$ that decrease with increasing futures horizons while $\hat{\alpha} < 0.5$ were estimated. For $p = 2$ however, the pattern is less consistent as no decreasing clustering of estimates $\hat{\alpha}$ across all 26 sets is observed as in the case for $p = 1$. The results are particularly interesting for the estimates implied by controlling for structural breaks by SIS in the case of $p = 2$ as we even observe a reversal of the pattern. The estimates $\hat{\alpha}$ having a tendency to increase with higher futures maturities. We remind however, that this peculiar results are unique to the case $p = 2$ when using individual financial instruments. The use of principal component analysis in order to reduce the number of variables from 26 to 5, while retaining over 60% of variation is a further robustness check. Indeed, the results are comparable to those implied by the EKT instrument sets for all three specifications for both cases $p = 1$ and $p = 2$.

However, this inconsistency for $p = 2$ does not come as a surprise: As Baumeister and Kilian (2016a) note in their evaluation, the risk premium estimates on crude oil futures with the help of typical financial variables may vary considerably, simply due to the inadequacy of a large number of commonly employed predictors in this field. A second explanation would be that fixing $p = 2$ does not correctly describe the underlying loss function. This explanation is reinforced as the inconsistent estimates are uniquely observed for $p = 2$ and the financial instrument sets. However, as Krüger and LeCrone (2019) point out, the J-tests have a high power against false models. As we observe very few rationality rejections, the evidence for a misspecified model while fixing $p = 2$ remains limited. Again, further research with high frequency and volatile data might be needed.

In order to evaluate, if market preferences evolve over time a rolling window estimation was used to identify two or three different sub-periods, which characterize our sample as a function of the value of set $p = 1$ or $p = 2$. Until around 1999, no clear tendency to under-predict or over-predict the oil spot price through futures pricing can be observed, as $\hat{\alpha}$ constantly oscillates around the symmetric value of 0.5. However, afterwards a continuous decrease in the value of $\hat{\alpha}$ emerges, pointing at a market preference to under-predict the spot price of oil through futures prices. In the case of $p = 1$ $\hat{\alpha}$ returns to the value of 0.5 at the end of the estimation sample. In the case of $p = 2$ $\hat{\alpha}$ returns to 0.5 in 2010 and continues to increase until the end of the sample period. This implies that from 2010 onwards, futures pricing systematically over-predicted the spot price of oil. It confirms the findings of Hamilton and Wu (2014), who show that the risk premium started to move from the long to the short side of futures contracts because of the increasing participation of index funds in futures markets. As Baumeister and Kilian (2016a) point out, it might have even become negative.

Similarly to the observations regarding the application of the MZ framework, we find that not controlling for outliers may result in very unstable GMM estimation results, not only with regard to forecast optimality testing, but more importantly with regard to the point estimates $\hat{\alpha}$. As we exemplarily showed, important "jumps" from high preferences to under-estimate to high preferences to highly over-estimate were observed between successive sample windows. Applying IIS was useful to remedy this instability.

In summary, based on the estimates $\hat{\alpha}$, regardless of the instruments used, strong evidence is provided for the existence of a time-varying risk premium that was negligible until the end of the 1990s. The risk premium became positive and reached its peak during the great oil price surge in the early 2000s, again descending to modest levels around 2015, but might have even become negative since 2005. The higher participation of Index Funds (Hamilton and Wu, 2014; Baumeister and Kilian, 2016a) in crude oil futures markets thus has reduced and even eliminated positive returns on futures contracts.

Chapter 9

Summary, conclusions & outlook

9.1 Summary & conclusions

This dissertation used time series methods to analyze four different aspects of crude oil research in economics. The real price of crude oil was at the center of our analyses. In chapter 5, one of the main empirical weaknesses in the early literature on crude oil, namely the assumption of exogenous crude oil prices with respect to the US and global business cycle has been reassessed with respect to interactions between macroeconomic performance and oil price shocks. The estimation of SVAR models allows for the disentanglement of the underlying causes of oil price shocks (see Kilian, 2009). The assumption is that unexpected oil price changes can be induced by global flow supply shocks, flow demand shocks resulting from global aggregate for industrial commodities including oil, and lastly oil-specific demand shocks. The latter shock also encompasses oil price changes because of expectations regarding future developments on global oil markets. By regressing monthly growth rates of the Dollar-Euro exchange rate, exports, imports, industrial production, unemployment and CPI inflation and quarterly series regarding GDP growth on the three different structural shock series, the reaction of the US and the German economies to unexpected oil price changes could be assessed.

The results confirm previous findings for the US economy: US macroeconomic aggregates react negatively to oil price increases, resulting from an increase in aggregate demand because of the global business cycle and to oil price increases because of oil-specific demand shocks. The German economy, in turn, is adversely affected by oil supply shocks. A positive effect on German macroeconomic aggregates can be observed in the analysis of oil price increases in response to an increase in global aggregate demand for industrial commodities. Moreover, no influence on the Dollar-Euro exchange rate can be observed. Keeping these main characteristics in mind, helps in the interpretation of previous empirical results that relied on oil price changes as explanatory variable, without differentiating for their underlying causes. Let's recall that historically, oil supply shocks have been much less prevalent and important in explaining unexpected oil price movements in comparison with aggregate demand shocks or oil-specific demand shocks. Therefore, in the framework of regressions using the oil price changes alone, aggregate demand shocks and oil-specific demand shocks are mostly the cause of the oil price changes.

For the US, such regressions solely on oil price changes result in mainly negative effects, whereas for Germany positive or no effects can be observed. Previous studies such as Kilian (2008a) and Blanchard and Gali (2007) that found no vulnerability of the German economy with regard to oil price shocks have thus to be reviewed based on our results. Indeed, with regard to the classical supply shocks attributed to political instabilities in the Middle East or other oil exporting countries, such as Venezuela or Nigeria (see e.g. Hamilton, 2011) an unexpected robustness of the US economy is observed, directly contradicting the findings of Hamilton (1983, 1985, 2011). While Germany on the other hand fits well into the perception of an economy vulnerable to oil supply disruptions. The neutral or positive effects previously observed in the case of Germany, were primarily due to strong aggregate demand that counteracts the negative effects due to increased oil prices.

A further promising extension comprises a review of previous studies that also found weak adverse or even positive effects of oil price shocks on other OECD economies such as the UK, France, Japan and Italy (see e.g. Kilian, 2008a; Guidi, 2010; Blanchard and Gali, 2007). In the same way as with Germany, weak negative effects on the economies, based on regressions relying uniquely on oil price changes, might be due to the weak effects resulting from oil-specific demand shocks or even positive effects that result from global aggregate demand shocks. We recall that both shocks primarily explain unexpected oil price movements in our sample.

Chapter 6 served the analysis of a specific oil supply shock in the form of sanctions on a crude oil exporter. The time-series regarding the Iranian oil production (see figure 6.2) clearly showed the negative impact on domestic crude oil production, when sanctions were implemented in 2011/12. In the same way, they clearly exhibit the immediate positive impact on the Iranian crude oil production after the removal of the sanction in 2015 with the implementation of the JCPOA. At first the forecast properties (as postulated by the MSE) of the recursively identified three-variable SVAR model and the sign-restricted identified four-variable SVAR model were compared under the assumption of knowing the impact of the implementation and removal of sanctions. It resulted in no model being superior to the other in both cases. Indeed, the four-variable model performed better when evaluating the negative supply shock after the implementation of the sanctions in 2011/12, whereas the three-variable model resulted in better forecasts upon the removal of the sanctions in 2015. A natural extension would be to estimate the three-variable sign identified SVAR model and compare the results. We note however, that as Kilian and Murphy (2012) point out in their estimation of the three-variable sign restricted model, the results are very similar, without the much higher computation requirements for the recursively identified model.

Based on the observed decline in Iranian oil production we were able to estimate the share of the observed structural crude oil supply shock series directly attributable to the Iranian oil sanctions. The structural relationships based on both SVAR specifications were required. This allowed the evaluation of the price increase of oil, attributed to the implementation of the Iranian oil boycott. Some surprising results are observed: The preferred four-variable SVAR model implied rising costs of a barrel of oil up to 0.4 \$, contrary to the three-variable SVAR model, which implied temporary increases of up to 8 \$ eight months after the implementation of the sanctions. A closer look at the evolution of crude oil prices as well as oil-specific demand shocks obtained by re-estimating of the three-variable SVAR with a sample up to December 2018, allowed a differentiated perception of the possible transmission of sanctions on crude oil prices.

Indeed, during the months preceding the formal imposition of both rounds of sanctions in 2011/12 and 2018, increasing oil prices can be observed. As the estimation of the structural shocks imply, this increase is primarily due to oil-specific demand shocks. The result is interpreted as an anticipation of future supply disruptions by market participants, that actually occur a few months later. Sanctions might have, therefore, an indirect flow supply shock effect through market expectations. In the SVAR models, these effects are not be included in the structural oil supply shock series, but in the oil-specific demand shock series.

These findings are considered to be important results of the analysis. They should be further investigated from different angles in order to provide a better understanding of structural shocks, implied by the global model of oil. It is possible that anticipated supply disruptions (not limited to sanctions) are not captured by the supply shock series but by the oil-specific demand shock series, thus underestimating the effects of flow supply shocks. When the supply disruptions are observed, their effects on prices remain limited as prices previously increased because of expectations. In this respect, a review of other exogenous events such as the second and third gulf wars in 1991 and 2003 is recommended. Comparable to the Iran sanctions, they resulted in the disruption of global oil supply, but were also preceded by a period with increased diplomatic tensions, thus giving market participants reason to anticipate future supply disruptions. A further case of a recent surprising oil supply shock, although short lived, was the drone attack on 14 September 2019 on the Saudi Arabian oil-processing facilities of Abqaiq–Khurais shutting down daily supply of around 5 mbd halving the Kingdoms daily output Reuters (2019b). Oil prices saw an immediate jump. WTI prices increased from 54.76 US$ to 63.1 US$, while Brent prices increased from 61.25 US$ to 68.42 US$. As the damages were quickly repaired, prices returned to pre-attack levels within two weeks.[1]

A closer review of events associated with or without real supply disruptions, would allow a better understanding of structural shocks resulting from the estimation of the global SVAR models of oil. In particular, we have reason to suspect that estimates of oil-specific demand shocks include anticipated supply disruptions that are observed and quantifiable at a later time. A recalibration of the SVAR models might thus impact the Hamilton-Kilian discussion concerning the relative importance of supply and demand disruptions. Attributing stronger effects to flow supply shocks would strengthening the original arguments described in Hamilton (1983). It would consequently also impact country analyses with regard to macroeconomic aggregates such as in chapter 5. We expect to estimate reduced effects attributed to oil-specific demand shocks and higher effects attributed to flow supply shocks. In the same manner, we would expect to find higher price effects of supply disruptions such as those implied by the Iran sanctions as seen in 5.

It has been confirmed in chapter 5 that unexpected oil price changes matter for economic performance. In this regard, chapters 7 and 8 highlight a different aspect of oil price research, namely oil price forecasting. Based on the three variables included in the global model of oil proposed by Kilian (2009), the possibilities to improve forecasts by using sparse selection methods were analyzed. Furthermore, different variable specifications as well the extension of the variable set were investigated, with the following main results: First, "long" VAR specifications that include many lags of the variables, such as required in the estimation procedure with respect to SVAR models because of impulse-response analysis, are dominated by more parsimonious VARs with respect to forecasting performance. Second, the variable transformations also matter. While again, in the case of the estimation of the above mentioned

[1]Based on the daily WTI and Brent price series as published by the EIA.

SVAR models, some variable specifications are required to control for impact and dynamic restrictions (see section 4.2.2), no such constraints exists in the case of the reduced form forecasts. Regularization improves forecasts for the longer forecast horizons up to 12 months for the VAR with variables transformed according to Kilian (2009) and the VAR in levels. For shorter horizons, the forecasts implied by the VAR in differences are also improved by applying the sparse estimators. Finally, we find no improvement by extending the variable set and applying regularization.

Chapter 8 complements the aforementioned forecast evaluation exercise by analyzing a financial contract that is commonly used as forecast of the spot price of oil, namely oil price futures. Using the framework developed by Elliott et al. (2005, 2008), a different view on the so-called risk premium on oil futures is offered. Previous studies relied on OLS estimates based on a wide range of regressors, resulting in very different and inconsistent estimates of the risk premium (see e.g. Baumeister and Kilian, 2016a).

The study results, in turn, consistently point at an asymmetric loss function: The estimates confirm a market preference to underestimate the spot price of crude oil ($\hat{\alpha} < 0.5$) through futures prices. This is accordance with the existence of a positive risk premium over the study sample, confirming the theory of normal backwardation. Furthermore, the tendency for $\hat{\alpha}$ to decrease with higher maturities is observed, implying increasing risk premia for longer horizons.

Our results should be taken into account in the current forecasting literature that is primarily focused on forecast combination approaches. Firstly, we propose to depart from the no-change forecast benchmark first used by Baumeister and Kilian (2015). A better understanding and comparison of various studies is only possible when incremental improvements are also included in subsequent studies. Secondly, we propose to combine our results regarding sparse estimation methods with future forecast combination evaluations. Thirdly we note, that we suspect a direct link from the asymmetry parameter estimate $\hat{\alpha}$ to the risk premium. An extension of our findings in order to estimate the risk premium based on $\hat{\alpha}$, could allow the construction of correct market expectations regarding the future spot price of oil. We recall that only futures prices and spot prices of crude oil, that are observed at a high frequency in real-time, are required within the EKT framework. In comparison and for example, the forecasts made by sparse VAR models such as in chapter 7, require data that is published with a lag of up to four months as in the case of the real activity index. A follow up study should thus enlighten the options available under this aspect.

9.2 Outlook

Finally we note, that this dissertation took the view on crude oil prices, as the majority of the literature, in the sense of market prices. As the Climate Accountability Institute (2019) estimates, between 1965-2017 the top 20 fossil fuel companies alone were responsible for 35% of global emissions, primarily through the combustion of their products. Toping the list were state-owned Saudi Aramco and investor-owned Chevron accounting for 4.38% and 3.2% of global emissions respectively. Here we refer to the United Nations Intergovernmental Panel on Climate Change (IPCC) for data and an overview of the research on global climate change that overwhelmingly attributes the increase in global greenhouse gas emissions to human economic activity.[2] So when analyzing the crude price of oil and it's impact on economies such as in chapter 5 future research should focus on the inclusion of the costs induced by global warming.

Furthermore, as various international institutions involved in global climate change policy management note in their latests outlooks regarding the emission goals agreed upon in the Paris Agreement, the phasing out from fossil fuels including crude oil will be crucial. Exemplary, the International Renewable Energy Agency (IRENA) notes in it's latest report, that daily crude oil consumption will have to decrease from around 95 mbd today to 60 mbd in 2030, 41 mbd in 2040 and finally 22 mbd in the year 2050 (International Renewable Energy Agency, 2019).

Needless to stress that such dramatic reduction in the size of the oil market would need transition management in order to avoid fights between oil producing countries, including the United States, for production quotas. Similarly, their reaction against a market competition by newcomers in the form of decentralized solar, wind and biomass producers would have to be carefully monitored. Such important reductions in crude oil demand would also mean, that a large part of the remaining hydrocarbon resources would remain underground for ever, as is already the case with most anthracite carbon reserves in Europe. Here we already saw the economic decline of regions that heavily relied on the extraction of coal. This transition was however mostly within national borders. As we saw in the case of crude oil, producing and consuming regions differ for the most part, requiring another level of international change management and policy instruments.

[2] `https://www.ipcc.ch/`.

Bibliography

Acharya, V. V., Lochstoer, L. A., and Ramadorai, T. (2013). Limits to arbitrage and hedging: Evidence from commodity markets. *Journal of Financial Economics*, 109(2):441 – 465.

Alquist, R. and Kilian, L. (2010). What do we learn from the price of crude oil futures? *Journal of Applied Econometrics*, 25(4):539–573.

Alquist, R., Kilian, L., and Vigfusson, R. (2013). Forecasting the price of oil. In Elliott, G., Granger, C., and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 2, chapter Chapter 8, pages 427–507. Elsevier.

Amano, R. and van Norden, S. (1995). Exchange rates and oil prices. *Bank of Canada*, Working Paper(95-8).

Amiti, M., Redding, S. J., and Weinstein, D. (2019). The impact of the 2018 trade war on us prices and welfare. Working Paper 25672, National Bureau of Economic Research.

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3):817–858.

Atkeson, A. and Kehoe, P. J. (1999). Models of energy use: Putty-putty versus putty-clay. *American Economic Review*, 89(4):1028–1043.

Backus, D. K. and Crucini, M. J. (2000). Oil prices and the terms of trade. *Journal of International Economics*, 50(1):185 – 213.

Barsky, R. and Kilian, L. (2004). Oil and the Macroeconomy Since the 1970s. CEPR Discussion Papers 4496.

Barsky, R. B. and Kilian, L. (2002). Do We Really Know that Oil Caused the Great Stagflation? A Monetary Alternative. In *NBER Macroeconomics Annual 2001, Volume 16*, NBER Chapters, pages 137–198. National Bureau of Economic Research, Inc.

Baumeister, C. and Kilian, L. (2014a). Real-Time Analysis of Oil Price Risks Using Forecast Scenarios. *IMF Economic Review*, 62(1):119–145.

Baumeister, C. and Kilian, L. (2014b). What central bankers need to know about forecasting oil prices. *International Economic Review*, 55(3):869–889.

Baumeister, C. and Kilian, L. (2015). Forecasting the Real Price of Oil in a Changing World: A Forecast Combination Approach. *Journal of Business & Economic Statistics*, 33(3):338–351.

Baumeister, C. and Kilian, L. (2016a). A General Approach to Recovering Market Expectations from Futures Prices with an Application to Crude Oil. Staff Working Papers 16-18, Bank of Canada.

Baumeister, C. and Kilian, L. (2016b). Forty years of oil price fluctuations: Why the price of oil may still surprise us. *Journal of Economic Perspectives*, 30(1):139–60.

Baumeister, C., Kilian, L., and Lee, T. K. (2014). Are there gains from pooling real-time oil price forecasts? *Energy Economics*, 46(S1):S33–S43.

Benassy-Quere, A., Mignon, V., and Penot, A. (2007). China and the Relationship between the Oil Price and the Dollar. *Energy Policy*, 35(11):5795–5805.

Bera, A. K. and Jarque, C. M. (1981). Efficient tests for normality, homoscedasticity and serial independence of regression residuals: Monte carlo evidence. *Economics Letters*, 7(4):313 – 318.

Bernanke, B. (1986). Alternative explanations of the money-income correlation. *Carnegie-Rochester Conference Series on Public Policy*, 25(1):49–99.

Bernanke, B., Gertler, M., and Watson, M. (1997). Systematic monetary policy and the effects of oil price shocks. *Brookings Papers on Economic Activity*, 28(1):91–157.

Bernanke, B. S. (1983). Irreversibility, Uncertainty, and Cyclical Investment. *The Quarterly Journal of Economics*, 98(1):85–106.

Bessembinder, H. (1992). Systematic risk, hedging pressure, and risk premiums in futures markets. *Review of Financial Studies*, 5(4):637–667.

Bessembinder, H. and Chan, K. (1992). Time-varying risk premia and forecastable returns in futures markets. *Journal of Financial Economics*, 32(2):169 – 193.

Blanchard, O. and Perotti, R. (2002). An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *The Quarterly Journal of Economics*, 117(4):1329–1368.

Blanchard, O. J. and Gali, J. (2007). The Macroeconomic Effects of Oil Shocks: Why are the 2000s So Different from the 1970s? NBER Working Papers 13368, National Bureau of Economic Research, Inc.

Blanchard, O. J. and Quah, D. (1989). The Dynamic Effects of Aggregate Demand and Supply Disturbances. *American Economic Review*, 79(4):655–673.

Bodenstein, M. and Guerrieri, L. (2011). Oil efficiency, demand, and prices: a tale of ups and downs. International Finance Discussion Papers 1031, Board of Governors of the Federal Reserve System (U.S.).

Broyden, C. G. (1970). The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations. *IMA Journal of Applied Mathematics*, 6(1):76–90.

Burbidge, J. and Harrison, A. (1984). Testing for the Effects of Oil-Price Rises Using Vector Autoregressions. *International Economic Review*, 25(2):459–84.

Canova, F. and Nicolo, G. D. (2002). Monetary disturbances matter for business fluctuations in the G-7. *Journal of Monetary Economics*, 49(6):1131–1159.

Capistrán, C. (2008). Bias in federal reserve inflation forecasts: Is the federal reserve irrational or just cautious? *Journal of Monetary Economics*, 55(8):1415 – 1427.

Carruth, A., Hooker, M., and Oswald, A. (1994). Unemployment, Oil Prices and the Real Interest Rate: Evidence from Canada and the UK. CEP Discussion Papers dp0188, Centre for Economic Performance, LSE.

Castle, J. L., Clements, M. P., and Hendry, D. F. (2016). An overview of forecasting facing breaks. *Journal of Business Cycle Research*, 12(1):3–23.

Castle, J. L., Doornik, J. A., Hendry, D. F., and Pretis, F. (2015). Detecting location shifts during model selection by step-indicator saturation. *Econometrics*, 3(2):240.

Chang, E. C. (1985). Returns to speculators and the theory of normal backwardation. *The Journal of Finance*, 40(1):193–208.

Cheng, F., Li, T., Wei, Y., and Fan, T. (2019). The VEC-NAR model for short-term forecasting of oil prices. *Energy Economics*, 78(C):656–667.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (2005). Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Journal of Political Economy*, 113(1):1–45.

Climate Accountability Institute (2019). Carbon majors: Uptade of top twenty companies 1965-2017. http://climateaccountability.org/pdf/CAI%20PressRelease%20Top20%20Oct19.pdf.

Cooley, T. and LeRoy, S. (1985). Atheoretical macroeconometrics: A critique. *Journal of Monetary Economics*, 16(3):283–308.

Cortright, D. and Lopez, G. (2000). *The sanctions decade : assessing UN strategies in the 1990s*.

Cragg, J. G. and Donald, S. (1993). Testing identifiability and specification in instrumental variable models. *Econometric Theory*, 9(02):222–240.

Croushore, D. (2012). Forecast bias in two dimensions. Working Papers 12-9, Federal Reserve Bank of Philadelphia.

Danthine, J.-P. (1978). Information, futures prices, and stabilizing speculation. *Journal of Economic Theory*, 17(1):79 – 98.

Darmstadter, J. (1971). *Energy in the World Economy*. The John Hopkins Press.

Davidson, R. and Mackinnon, J. (2009). *Econometric Theory and Methods*. Oxford University Press.

De Roon, F. A., Nijman, T. E., and Veld, C. (2000). Hedging pressure effects in futures markets. *The Journal of Finance*, 55(3):1437–1456.

Degiannakis, S. and Filis, G. (2018). Forecasting oil prices: High-frequency financial data are indeed useful. *Energy Economics*, 76(C):388–402.

Dvir, E. and Rogoff, K. S. (2009). Three epochs of oil. Working Paper 14927, National Bureau of Economic Research.

Eaton, J. and Engers, M. (1992). Sanctions. *Journal of Political Economy*, 100(5):899–928.

Eaton, J. and Engers, M. (1999). Sanctions: Some simple analytics. *American Economic Review*, 89(2):409–414.

Edelstein, P. and Kilian, L. (2007a). Retail Energy Prices and Consumer Expenditures. CEPR Discussion Papers 6255, C.E.P.R. Discussion Papers.

Edelstein, P. and Kilian, L. (2007b). The Response of Business Fixed Investment to Changes in Energy Prices: A Test of Some Hypotheses about the Transmission of Energy Price Shocks. *The B.E. Journal of Macroeconomics*, 7(1):1–41.

Edelstein, P. and Kilian, L. (2009). How sensitive are consumer expenditures to retail energy prices? *Journal of Monetary Economics*, 56(6):766–779.

Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R. (2004). Least angle regression. *Ann. Statist.*, 32(2):407–499.

Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Number 57 in Monographs on Statistics and Applied Probability. Chapman & Hall/CRC, Boca Raton, Florida, USA.

Elliott, G., Komunjer, I., and Timmermann, A. (2005). Estimation and testing of forecast rationality under flexible loss. *Review of Economic Studies*, 72(4):1107–1125.

Elliott, G., Komunjer, I., and Timmermann, A. (2008). Biases in macroeconomic forecasts: Irrationality or asymmetric loss? *Journal of the European Economic Association*, 6(1):122–157.

Elliott, G., Rothenberg, T. J., and Stock, J. H. (1996). Efficient tests for an autoregressive unit root. *Econometrica*, 64(4):813–836.

Energy Information Administration (2017). World oil transit chokepoints. `https://www.eia.gov/beta/international/regions-topics.php?RegionTopicID=WOTC`.

Energy Information Administration (2019a). Country analysis executive summary: Iran. `https://www.eia.gov/international/content/analysis/countries_long/Iran/pdf/iran_exe.pdf`.

Energy Information Administration (2019b). What are petroleum products, and what is petroleum used for? `https://www.eia.gov/tools/faqs/faq.php?id=41&t=6`.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–1007.

Engle, R. F. and Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, 55(2):251–276.

European Council (2019). Joint comprehensive plan of action and restrictive measures. https://www.consilium.europa.eu/de/policies/sanctions/iran/jcpoa-restrictive-measures/.

Fama, E. F. and French, K. R. (1987). Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage. *The Journal of Business*, 60(1):55–73.

Fama, E. F. and French, K. R. (1988). Dividend yields and expected stock returns. *Journal of Financial Economics*, 22(1):3 – 25.

Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 96(456):1348–1360.

Fattouh, B., Kilian, L., and Mahadeva, L. (2013). The role of speculation in oil markets: What have we learned so far? *Energy Journal*, 34:7–33.

Faust, J. (1998). The robustness of identified VAR conclusions about money. *Carnegie-Rochester Conference Series on Public Policy*, 49(1):207–244.

Ferderer, J. P. (1996). Oil price volatility and the macroeconomy. *Journal of Macroeconomics*, 18(1):1 – 26.

Finn, M. G. (2000). Perfect competition and the effects of energy price increases on economic activity. *Journal of Money, Credit and Banking*, 32(3):400–416.

Fry, R. and Pagan, A. (2005). Some Issues In Using Vars For Macroeconometric Research. CAMA Working Papers 2005-19, Centre for Applied Macroeconomic Analysis, Crawford School of Public Policy, The Australian National University.

Fry, R. and Pagan, A. (2011). Sign restrictions in structural vector autoregressions: A critical review. *Journal of Economic Literature*, 49(4):938–960.

Funk, C. (2018). Forecasting the real price of oil - time-variation and forecast combination. *Energy Economics*, 76:288–302.

Ganser, D. (2015). *Europa im Erdölrausch*. Orell Füssli.

Garratt, A., Vahey, S. P., and Zhang, Y. (2019). Real-time forecast combinations for the oil price. *Journal of Applied Econometrics*, 34(3):456–462.

Gisser, M. and Goodwin, T. H. (1986). Crude oil and the macroeconomy: Tests of some popular notions. *Journal of Money, Credit and Banking*, 18(1):95–103.

Gorton, G. B., Hayashi, F., and Rouwenhorst, K. G. (2012). The fundamentals of commodity futures returns. *Review of Finance*.

Graf, R. (2014). *Öl und Souveränität - Petroknowledge und Energiepolitik in den USA und Westeuropa in den 1970er Jahren*. De gruyter Oldenbourg.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438.

Grossman, S. J. (1977). The existence of futures markets, noisy rational expectations and informational externalities. *The Review of Economic Studies*, 44(3):431–449.

Guidi, F. (2010). The Economic Effects of Oil Prices Shocks on the UK Manufacturing and Services Sectors. *The IUP Journal of Applied Economics*, 0(4):5–34.

Guo, M., Lu, L., Sheng, L., and Yu, M. (2018). The day after tomorrow: Evaluating the burden of trump's trade war. *Asian Economic Papers*, 17(1):101–120.

Hall, R. (1991). Labor demand, labor supply, and employment volatility. *NBER Macroeconomics Annual*, 6:17.

Hamilton, J. (1988). A neoclassical model of unemployment and the business cycle. *Journal of Political Economy*, 96(3):593–617.

Hamilton, J. D. (1983). Oil and the Macroeconomy since World War II. *Journal of Political Economy*, 91(2):228–48.

Hamilton, J. D. (1985). Historical causes of postwar oil shocks and recessions. *The Energy Journal*, 6(1):97–116.

Hamilton, J. D. (1996). This is what happened to the oil price-macroeconomy relationship. *Journal of Monetary Economics*, 38(2):215 – 220.

Hamilton, J. D. (2001). A parametric approach to flexible nonlinear inference. *Econometrica*, 69(3):537–573.

Hamilton, J. D. (2003). What is an oil shock? *Journal of Econometrics*, 113(2):363–398.

Hamilton, J. D. (2009a). Causes and Consequences of the Oil Shock of 2007-08. *Brookings Papers on Economic Activity*, 40(1 (Spring):215–283.

Hamilton, J. D. (2009b). Understanding crude oil prices. *Energy Journal*, 30.

Hamilton, J. D. (2011). Historical Oil Shocks. NBER Working Papers 16790, National Bureau of Economic Research, Inc.

Hamilton, J. D. (2016). *Oil and the Macroeconomy*, pages 1–7. Palgrave Macmillan UK, London.

Hamilton, J. D. (2018). Measuring global economic activity. *https://econweb.ucsd.edu/ jhamilto/REA.pdf*.

Hamilton, J. D. and Herrera, A. M. (2004). Oil Shocks and Aggregate Macroeconomic Behavior: The Role of Monetary Policy: Comment. *Journal of Money, Credit and Banking*, 36(2):265–286.

Hamilton, J. D. and Wu, J. C. (2014). Risk premia in crude oil futures prices. *Journal of International Money and Finance*, 42:9 – 37. Understanding International Commodity Price Fluctuations.

Hansen, L., Heaton, J., and Yaron, A. (1996). Finite-sample properties of some alternative gmm estimators. *Journal of Business & Economic Statistics*, 14(3):262–80.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4):1029–1054.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The elements of statistical learning: data mining, inference and prediction*. Springer, 2 edition.

Hausman, J. and Newey, W. (1995). Nonparametric estimation of exact consumers surplus and deadweight loss. *Econometrica*, 63(6):1445–76.

Hendry, D. F. and Doornik, J. A. (2014). *Empirical model discovery and theory evaluation : automatic selection methods in econometrics*. Arne Ryde memorial lectures. The MIT Press, Cambridge, Massachusetts.

Herrera, A. M. and Pesavento, E. (2009). Oil price shocks, systematic monetary policy and the great moderation. *Macroeconomic Dynamics*, 13(1):107–137.

Hoffman, D. and Rasche, R. (1996). Assessing forecast performance in a cointegrated system. *Journal of Applied Econometrics*, 11(5):495–517.

Hong, H. and Yogo, M. (2012). What does futures market interest tell us about the macroeconomy and asset prices? *Journal of Financial Economics*, 105(3):473 – 490.

Hooker, M. A. (1996). What happened to the oil price-macroeconomy relationship? *Journal of Monetary Economics*, 38(2):195–213.

Hooker, M. A. (2002). Are oil shocks inflationary? asymmetric and nonlinear specifications versus changes in regime. *Journal of Money, Credit and Banking*, 34(2):540–61.

Hufbauer, G. C., Schott, J. J., and Elliott, K. A. (2009). *Economic Sanctions Reconsidered 3rd edition (hardcover + CD)*. Number 4082 in Peterson Institute Press: All Books. Peterson Institute for International Economics.

Hults, D. R. (2012). Petróleos de Venezuela, S.A. (PDVSA): from independence to subservience. In Victor, D. G., Hults, D. R., and Thurber, M., editors, *Oil and Governance: State-Owned Enterprises and the World Energy Supply*, pages 418–477. Cambridge University Press.

Hyndman, R. J. and Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4):679 – 688.

International Renewable Energy Agency (2019). *Global energy transformation: A roadmap to 2050 (2019 edition)*. Abu Dhabi.

Jarque, C. M. and Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3):255 – 259.

Jarque, C. M. and Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review / Revue Internationale de Statistique*, 55(2):163–172.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2):231 – 254.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica*, 59(6):1551–80.

Jolliffe, I. (2002). *Principal component analysis*. Springer Verlag, New York.

Katzman, K. (2019). Iran sanctions. Technical report, Congressional Research Service.

Keane, M. and Runkle, D. E. (1990). Testing the rationality of price forecasts: New evidence from panel data. *American Economic Review*, 80(4):714–35.

Kellog, R. (2011). Short-run responses of oil production to spot and futures oil prices. mimeo.

Keynes, J. M. (1930). *A Treatise on Money*, volume 2. Macmillan, London.

Kilian, L. (2008a). A Comparison of the Effects of Exogenous Oil Supply Shocks on Output and Inflation in the G7 Countries. *Journal of the European Economic Association*, 6(1):78–121.

Kilian, L. (2008b). Exogenous Oil Supply Shocks: How Big Are They and How Much Do They Matter for the U.S. Economy? *The Review of Economics and Statistics*, 90(2):216–240.

Kilian, L. (2008c). The Economic Effects of Energy Price Shocks. *Journal of Economic Literature*, 46(4):871–909.

Kilian, L. (2009). Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market. *American Economic Review*, 99(3):1053–69.

Kilian, L. (2013). Structural vector autoregressions. In Hashimzade, N. and Thornton, M. A., editors, *Handbook of Research Methods and Applications in Empirical Macroeconomics*, chapter 22, pages 515–554. Edward Elgar Publishing, Inc., Cheltenham, UK.

Kilian, L. (2014). Oil Price Shocks: Causes and Consequences. *Annual Review of Resource Economics*, 6(1):133–154.

Kilian, L. (2017). The Impact of the Fracking Boom on Arab Oil Producers. *The Energy Journal*, 38(6):137–160.

Kilian, L. (2018). Measuring global real economic activity: Do recent critiques hold up to scrutiny? *https://sites.google.com/site/lkilian2019/home*.

Kilian, L. and Lee, T. K. (2014). Quantifying the speculative component in the real price of oil: The role of global oil inventories. *Journal of International Money and Finance*, 42(C):71–87.

Kilian, L. and Lütkepohl, H. (2017). *Structural Vector Autoregressive Analysis*. Themes in Modern Econometrics. Cambridge University Press.

Kilian, L. and Murphy, D. (2012). Why agnostic sign restrictions are not enough: Understanding the dynamics of oil market var models. *Journal of the European Economic Association*, 10(5):1166–1188.

Kilian, L. and Murphy, D. P. (2014). The Role Of Inventories And Speculative Trading In The Global Market For Crude Oil. *Journal of Applied Econometrics*, 29(3):454–478.

Kilian, L. and Park, C. (2009). The impact of oil price shocks on the us stock market. *International Economic Review*, 50(4):1267–1287.

Kilian, L. and Zhou, X. (2018). Modeling fluctuations in the global demand for commodities. *Journal of International Money and Finance*, 88(C):54–78.

Kim, I.-M. and Loungani, P. (1992). The role of energy in real business cycle models. *Journal of Monetary Economics*, 29(2):173 – 189.

Krüger, J. and LeCrone, J. (2019). Forecast evaluation under asymmetric loss: A monte carlo analysis of the ekt method. *Oxford Bulletin of Economics and Statistics*, 81(2):437–455.

Krugman, P. (1983a). Oil and the dollar. In Bhandari, J. and Putnam, B., editors, *Economic Interdependence and Flexible Exchange Rates*, pages 179–190. MIT Press.

Krugman, P. (1983b). Oil shocks and exchange rate dynamics. In Frankel, J., editor, *Exchange Rates and International Macroeconomics*, pages 259–284. University of Chicago Press.

Kwiatkowski, D., Phillips, P., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. how sure are we that economic time series have unit root? *Journal of Econometrics*, 54:159–178.

Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer, Berlin.

Mahdavi, P. (2012). Oil, monarhcy, revolution, and theocracy: a study on the national iranian oil company (nioc). In Victor, D. G., Hults, D. R., and Thurber, M., editors, *Oil and Governance: State-Owned Enterprises and the World Energy Supply*, chapter 6, pages 234–279. Cambridge University Press.

Mayer, F. (1966). *Erdöl Weltatlas*. Gerog Westermann Verlag Braunschweig.

Mincer, J. A. and Zarnowitz, V. (1969). The Evaluation of Economic Forecasts. In *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, NBER Chapters, pages 3–46. National Bureau of Economic Research, Inc.

Mork, K. A. (1989). Oil and Macroeconomy When Prices Go Up and Down: An Extension of Hamilton's Results. *Journal of Political Economy*, 97(3):740–44.

Murphy, K. (2012). *Machine Learning: A Probabilistic Perspective*. Cambridge (Mass.): MIT Press.

Nakov, A. and Pescatori, A. (2010). Monetary policy trade-offs with a dominant oil producer. *Journal of Money, Credit and Banking*, 42(1):1–32.

Neuenkirch, M. and Neumeier, F. (2015). The impact of un and us economic sanctions on gdp growth. *European Journal of Political Economy*, 40(PA):110–125.

New York Times (2017). The iran crisis explained. `https://www.nytimes.com/2019/06/17/us/politics/iran-nuclear-deal-uranium.html`.

Newey, W. K. and West, K. D. (1987). A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, 55(3):703–708.

Nicholson, W. B., Matteson, D. S., and Bien, J. (2016). *High Dimensional Forecasting via Interpretable Vector Autoregression*. arXiv:1412.5250v2.

Nicholson, W. B., Matteson, D. S., and Bien, J. (2017). VARX-L: Structured regularization for large vector autoregressions with exogenous variables. *International Journal of Forecasting*, 33(3):627–651.

Pagano, P. and Pisani, M. (2009). Risk-adjusted forecasts of oil prices. Working Paper Series 0999, European Central Bank.

Patterson, R. (2013). EU Sanctions on Iran: The European Political Context. *Middle East Policy*.

Patton, A. J. and Timmermann, A. (2012). Forecast rationality tests based on multi-horizon bounds. *Journal of Business & Economic Statistics*, 30(1):1–17.

Peersman, G. and van Robays, I. (2009). Oil and the Euro area economy. *Economic Policy*, 24(60):603–651.

Peksen, D. (2009). Better or worse? the effect of economic sanctions on human rights. *Journal of Peace Research*, 46(1):59–77.

Peksen, D. and Drury, A. C. (2010). Coercive or corrosive: The negative impact of economic sanctions on democracy. *International Interactions*, 36(3):240–264.

Pfaff, B. (2008). VAR, SVAR and SVEC Models: Implementation Within R Package vars. *Journal of Statistical Software*, 27(i04).

Pretis, F., Mann, M. L., and Kaufmann, R. K. (2015). Testing competing models of the temperature hiatus: assessing the effects of conditioning variables and temporal uncertainties through sample-wide break detection. *Climatic Change*, 131(4):705–718.

Ramey, V. (1991). Nonconvex costs and the behavior of inventories. *Journal of Political Economy*, 99(2):306–34.

Reuters (2019a). Hit by sanctions and rising tensions, Iran's oil exports slide in July. `https://www.reuters.com/article/us-oil-iran-exports/hit-by-sanctions-and-rising-tensions-irans-oil-exports-slide-in-july-idUSKCN1UP1UD`.

Reuters (2019b). Oil jumps nearly 15% in record trading after attack on Saudi facilities. `https://www.reuters.com/article/us-global-oil/oil-prices-surge-15-after-attack-on-saudi-facilities-hits-global-supply-idUSKBN1W00UG`.

Romer, C. D. and Romer, D. H. (2000). Federal Reserve Information and the Behavior of Interest Rates. *American Economic Review*, 90(3):429–457.

Rossi, B. and Sekhposyan, T. (2016). Forecast Rationality Tests in the Presence of Instabilities, with Applications to Federal Reserve and Survey Forecasts. *Journal of Applied Econometrics*, 31(3):507–532. jae.2440.

Rotemberg, J. J. and Woodford, M. (1996). Imperfect Competition and the Effects of Energy Price Increases on Economic Activity. *Journal of Money, Credit and Banking*, 28(4 Part 1):549–577.

Rubio-Ramirez, J. F., Waggoner, D., and Zha, T. (2010). Structural vector autoregressions: Theory of identification and algorithms for inference. *Review of Economic Studies*, 77(2):665–696.

Sadorsky, P. (2002). Time-varying risk premiums in petroleum futures prices. *Energy Economics*, 24(6):539 – 556.

Santini, D. J. (1985). The energy-squeeze model: energy price dynamics in US business cycles. *International Journal of Energy Systems*, 5(1):18–25.

Santos, C., Hendry, D. F., and Johansen, S. (2008). Automatic selection of indicators in a fully saturated regression. *Computational Statistics*, 23(2):317–335.

Sims, C., Stock, J., and Watson, M. (1990). Inference in linear time series models with some unit roots. *Econometrica*, 58(1):113–44.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, 48(1):1–48.

Sims, C. A. (1986). Are forecasting models usable for policy analysis? *Quarterly Review*, (Win):2–16.

Sims, C. A. (2002). The Role of Models and Probabilities in the Monetary Policy Process. *Brookings Papers on Economic Activity*, 2002(2):1–40.

Sinclair, T. M., Joutz, F., and Stekler, H. (2010). Can the FED predict the state of the economy? *Economics Letters*, 108(1):28 – 32.

Singleton, K. J. (2014). Investor flows and the 2008 boom/bust in oil prices. *Management Science*, 60(2):300–318.

Smil, V. (2000). Energy in the twentieth century: Resources, conversions, costs, uses, and consequences. *Annual Review of Energy and the Environment*, 25(1):21–51.

Stevens, P. (2012). Saudi Aramco: the Jewel in the Crown. In Victor, D. G., Hults, D. R., and Thurber, M., editors, *Oil and Governance: State-Owned Enterprises and the World Energy Supply*, chapter 5, pages 173–233. Cambridge University Press.

Stillwagon, J. R. (2016). Non-linear exchange rate relationships: An automated model selection approach with indicator saturation. *The North American Journal of Economics and Finance*, 37:84 – 109.

Stock, J. and Watson, M. (2003). Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature*, 41(3):788–829.

Stopford, M. (1997). *Maritime Economics*. Routledge.

Sucarrat, G., Pretis, F., and Reade, J. (2016). *gets: General-to-Specific (GETS) Modelling and Indicator Saturation Methods*. R package version 0.7.

Tatom, J. A. (1988). Are the macroeconomic effects of oil-price changes symmetric? *Carnegie-Rochester Conference Series on Public Policy*, 28:325 – 368.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288.

Uhlig, H. (2005). What are the effects of monetary policy on output? Results from an agnostic identification procedure. *Journal of Monetary Economics*, 52(2):381–419.

Van de Graaf, T. (2013). The Oil Weapon Reversed? Sanctions Against Iran and US-EU Structural Power. *Middle East Policy*, 20.

Wang, Y., Liu, L., and Wu, C. (2017). Forecasting the real prices of crude oil using forecast combinations over time-varying parameter models. *Energy Economics*, 66(C):337–348.

Yergin, D. (1991). *The Prize - The Epic Quest for Oil, Money, and Power*. Simon & Schuster Inc.

Zhang, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. *Ann. Statist.*, 38(2):894–942.

Zhang, Y., Ma, F., Shi, B., and Huang, D. (2018). Forecasting the prices of crude oil: An iterated combination approach. *Energy Economics*, 70(C):472–483.

Zhang, Y., Ma, F., and Wang, Y. (2019). Forecasting crude oil prices with a large set of predictors: Can lasso select powerful predictors? *Journal of Empirical Finance*, 54:97 – 117.

Zhou, S. (1995). The response of real exchange rates to various economic shocks. *Southern Economic Journal*, 61(4):936–954.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 67(2):301–320.

Appendix

Appendix: Rolling window results chapter 7

The figures 1, 2 and 3 in this appendix show the results with a rolling window design of the forecast experiment. Used is a rolling window of 240 months for the estimation sample.

Summarizing these results we can conclude that the overall pattern of results is quite similar when using the rolling window instead of the expanding window. However, we generally find larger forecast errors and larger final MSE values at the end of the evaluation period. Using the sparse VAR methods results in smaller improvements for the shorter forecast horizons and no visible improvements and even deteriorations at the longer horizons. In particular, for the VAR in levels we find a tremendous deterioration of LASSO and ENET at longer horizons in the aftermath of the financial crisis which does not occur when using the expanding window. Invoking the IIS procedure does only occasionally lead to improvements of the forecast performance and generally causes the forecast errors to be larger.

Overall and irrespective of the variable transformations we find a better forecast performance at the end of the sample period with the expanding window instead of the rolling window. This outcome may be a cause of the larger sample size available when using the expanding window. A major contribution can attributed to the financial crisis which is comprised in each of the rolling window estimation samples until the end of the evaluation period with a larger weight than in the expanding windows.

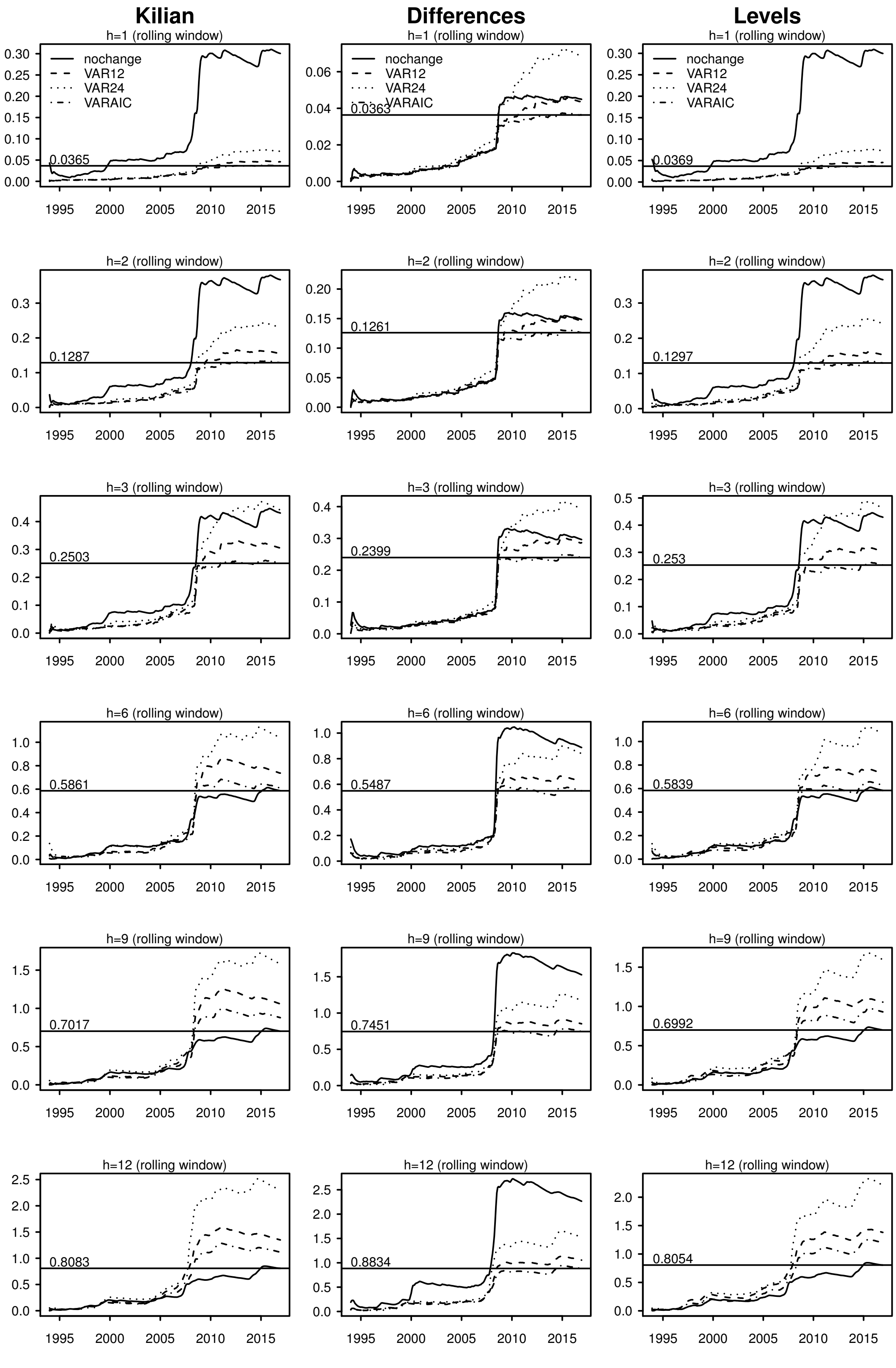

Figure 1: Benchmark selection (rolling window).

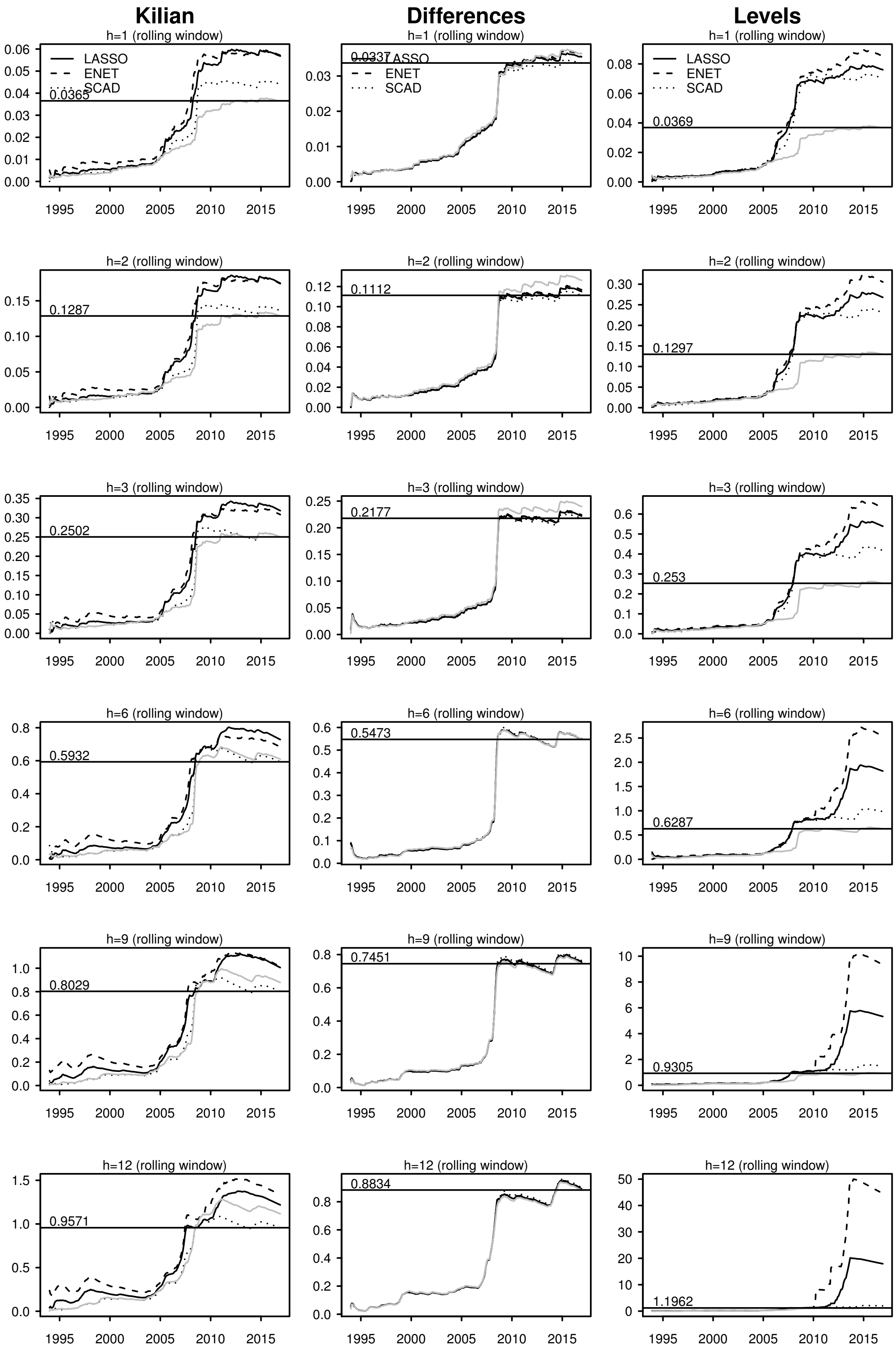

Figure 2: Evaluation of the sparse VARs (rolling window).

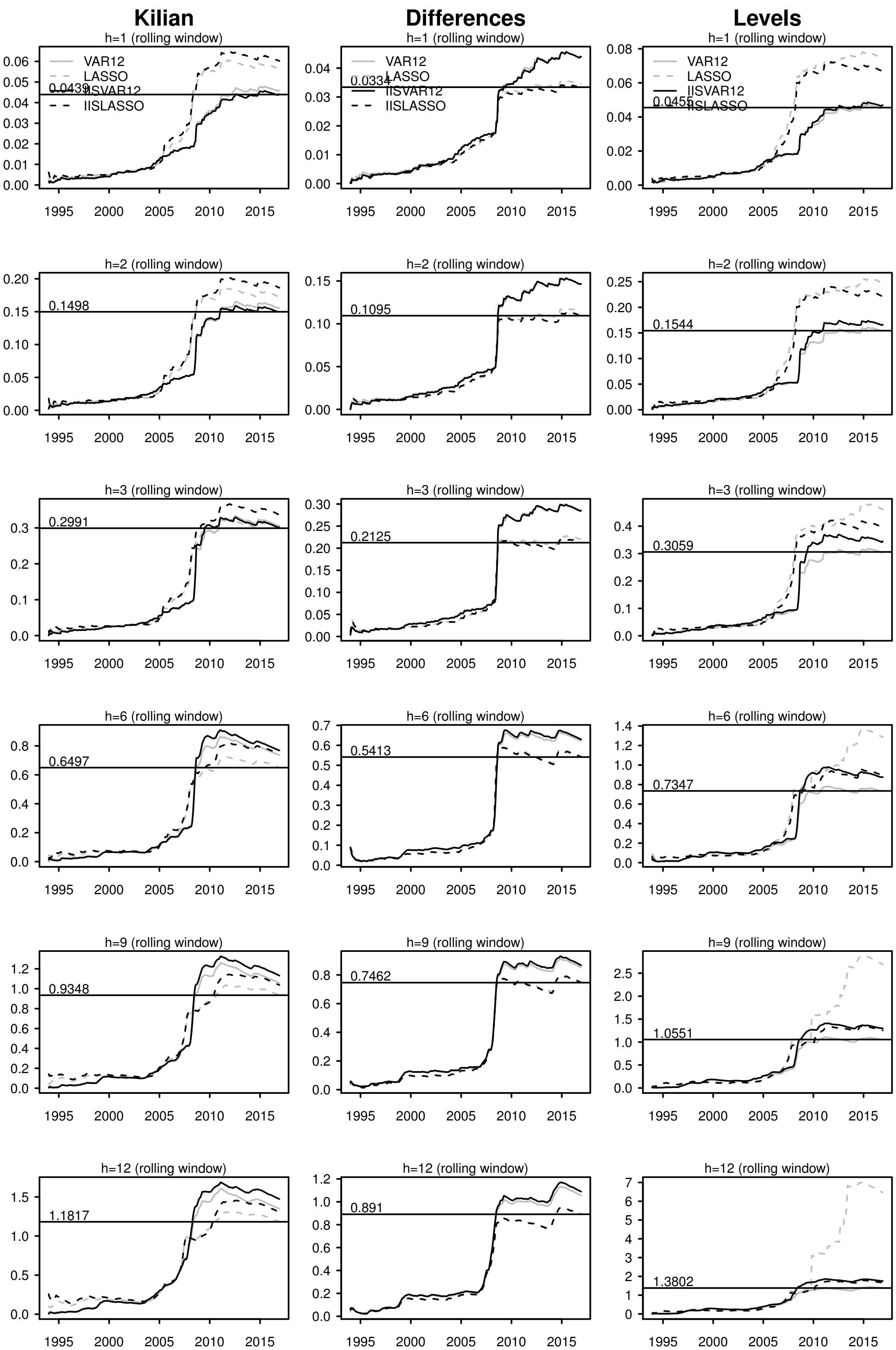

Figure 3: Evaluation with IIS (rolling window).

Appendix: Data sources for financial variables employed in chapter 8

Instrument set	Data sources	Notes
1	Homepage Kenneth. R. French http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html	Median market return
2	FRED database with code CPIAUCSL	Inflation - expected inflation
3	FRED database with code GS1M	1-month T-Bill - average interest rate over the past 12 months
4	FRED database with code GS3M	
5	FRED database with code GS20, GS3M	
6	FRED database with code BAA, GS20	
7	FRED database with code CPI	Fitted residuals from an ARIMA(3,1,0) model regarding industrial production
8	Homepage Kenneth. R. French http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html	Median dividend yield
9	FRED database with code SP500	
10	http://www.multpl.com/s-p-500-dividend-yield/table?f=m	
11	FRED database with code AAA, BAA	
12	FRED database with code GS3M	
13	Homepage Kenneth. R. French http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html FRED database with code DGS1MO	Median market return - 1-month T-Bill
14	FRED database with code MCUMFN	
15-17	FRED database with codes GS1, GS2, GS5, GS10	
18	https://data.oecd.org/leadind/composite-leading-indicator-cli.htm	
19	FRED database with code AAA, GS1M	
20-25	Futures prices from the Thomson Reuters database with codes NCLS00, NCLS02, NCLS03, NCLS04, NCLS05, NCLS06. WTI spot price from the FRED database with the series code CRUDOIL	$(\frac{f_{t+h}}{y_{t+h}})^h$ f_{t+h}: price of h^{th} futures in t y_{t+h}: spot price in t
26	US Energy Information Agency	

Affirmation

I hereby declare that the dissertation entitled

"Time Series Analyses of Global Oil Prices: Shocks, Effects and Predictability"

is my own work. I have only used the sources indicated and have not made unauthorized use of services of a third party. Where the work of others has been quoted or reproduced, the source is always given. I have not presented this thesis or parts thereof to an university as part of an examination or degree.

....

Place, Date Signature